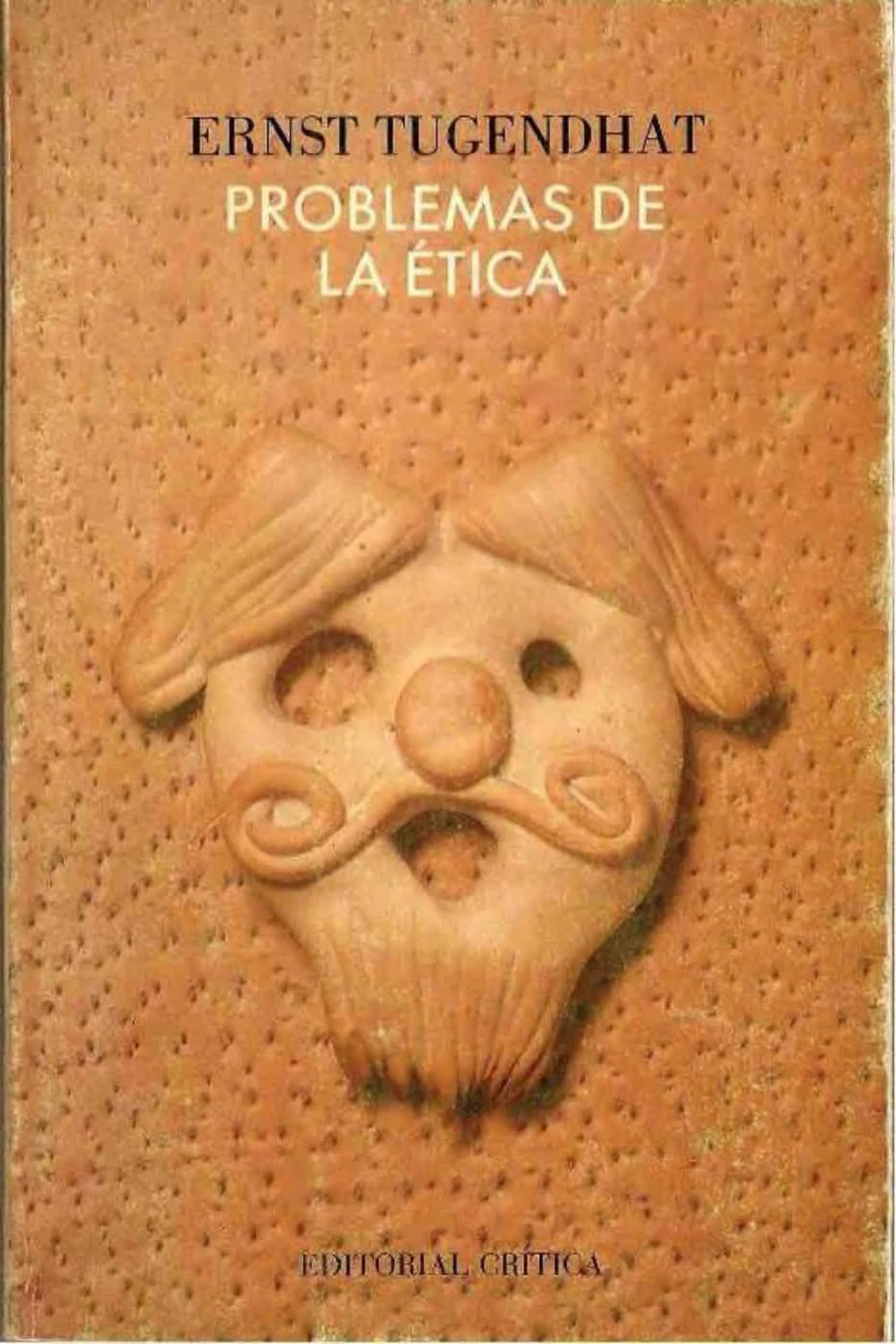
ERNST TUGENDHAT

PROBLEMAS DE LA ÉTICA



EDITORIAL CRÍTICA

ERNST TUGENDHAT (Brno, Checoslovaquia, 1930) se licenció en filología clásica en la Universidad de Stanford e hizo el doctorado en filosofía en Friburgo. Actualmente es catedrático en la Universidad Libre de Berlín Occidental. Su obra tuvo al principio un enfoque básicamente epistemológico, centrado en el análisis minucioso de ciertos conceptos como el de «verdad», así como en el estudio y balance de la filosofía analítica. En los últimos años los temas éticos polarizan toda su actividad intelectual, no limitándose en ellos a las cuestiones más teóricas y abstractas: junto al problema de los fundamentos de la ética, objeto primordial de *Problemas de la ética*, Tugendhat polemiza con el tratamiento político de ciertos asuntos prácticos, concretamente el problema del armamentismo nuclear.

PROBLEMAS DE LA ÉTICA

CRÍTICA/FILOSOFÍA
Directora: VICTORIA CAMPS

ERNST TUGENDHAT

PROBLEMAS DE LA ETICA

Traducción castellana de
JORGE VIGIL

EDITORIAL CRÍTICA
Grupo editorial Grijalbo
BARCELONA

Título original:
PROBLEME DER ETHIK (cc)(i)($)(=)

Cubierta: Enric Satué sobre un trabajo artesanal, en pan, de Eduardo Crespo
© 1984: Philipp Reclam jun., Stuttgart
© 1988 de la traducción castellana para España y América: Editorial Crítica, S. A., Aragó, 385, 08013 Barcelona
ISBN: 84-7423-344-5
Depósito legal: B. 1.802-1988
Impreso en España
1988. — NOVAGRAFIK, Puigcerdà, 127, 08019 Barcelona

PRÓLOGO

¿Qué significa decir de una acción que es *mala* (mala sin más, no para este o aquel fin)? ¿Qué queremos dar a entender cuando decimos que se *debe* o *tiene que* obrar así, o que *no se puede* obrar así (de forma categórica, y no con respecto a un determinado propósito)? ¿Cómo pueden fundamentarse expresiones de este tipo? ¿Acaso no pueden fundamentarse? ¿Qué sentido tiene hablar aquí de fundamentación? Éstas son las preguntas centrales de la ética o filosofía moral.

Estas preguntas tan fundamentales para la convivencia humana han sido las preguntas centrales de la filosofía desde sus mismos orígenes; puede incluso decirse que con ellas ha nacido la filosofía. En los últimos treinta años se ha registrado de nuevo un interés especialmente intenso por ellas, patente sobre todo en Alemania durante estos quince últimos años. Sin embargo, hasta la fecha no se han ofrecido respuestas demasiado convincentes. Sin duda esto se debe en parte a su misma naturaleza, pues es característico de las preguntas filosóficas, que tienen que ser formuladas siempre de nuevo, el no hallar respuestas definitivas. No obstante, creo que una gran parte de la abundante literatura filosófico-moral de nuestra época resulta comparativamente tan insatisfactoria porque en general (y

aquí me incluyo a mí mismo) hasta ahora se ha abordado el problema con la falsa suposición de que hay que dar una respuesta breve y sencilla —ya sea escéptica o positiva— a estas preguntas. Pienso que esta problemática no puede ser tratada cabalmente mientras se pretenda poder resolverla de una vez por todas. Se trata de un objeto genuino de investigación, y por cierto una investigación en la que tienen que unirse métodos específicamente filosóficos —analítico-conceptuales— y empíricos.

Esta concepción choca con otra muy difundida en la filosofía moral según la cual hay que separar estrictamente de los hechos empíricos la pregunta por la «validez» de las normas morales. Quien no observa esta distinción entre «es» y «debe» incurre en una «falacia naturalista» (*naturalistic fallacy*): aquello que *debemos* hacer no puede fundarse —se dice— en ningún tipo de constataciones empíricas sobre lo que *es*. A primera vista, esto parece muy plausible. Sin embargo, tras una reflexión minuciosa surgen algunas dificultades.

En primer lugar, si lo debido no puede fundamentarse empíricamente, ¿cómo se puede fundamentar? La respuesta sólo puede ser: apriorísticamente. Así lo ha hecho, por ejemplo, Kant. Pero para ello se necesita además el concepto de un *a priori* sustancial y no sólo analítico; Kant tenía un *a priori* semejante, pero nosotros no. Un concepto así implica que no somos sólo seres de este mundo terrenal, sino además miembros de otro mundo precedente y superior. El concepto kantiano de *a priori* sintético es —según puede verse hoy— un desafortunado intento de secularización de la transcendencia en sentido religioso.

La segunda dificultad es aun más fundamental: quienes consideran obvio que lo que debe ser no se

puede reducir a lo que es, suponen que comprendemos sin más qué significa el «deber»/«tener que» en este sentido absoluto (moral). En realidad se trata de un concepto hasta ahora no explicado, y lo seguirá siendo siempre que se le dé el estatus de una «validez» apriorística. En vez de invocar una evidencia, la de que lo que debe es quizás «otra cosa» que lo que es, se trata más bien de clarificar qué entendemos por «debe» en el sentido moral de la palabra. Como opino que se han ensayado todas las posibilidades «simples» de comprender este uso del «debe», habiendo fracasado todas ellas, he llegado a la conclusión de que no se puede alcanzar la explicación analítica de este significado sin la cooperación del filósofo con el investigador empírico de la socialización (véase p. 154).

Las «Tres lecciones sobre problemas de la ética» y las «Retractaciones» a ellas referidas que aquí publico constituyen mis dos últimos y provisionales pasos de una serie de varios intentos —que siempre he considerado fracasados— de comprensión de los enunciados morales.

Un primer intento lo hice en dos lecciones sobre ética profesadas en Heidelberg durante los semestres de invierno de los años 1967-1968 y 1973-1974. En el séptimo capítulo de mis *Vorlesungen zur Einführung in die sprachanalytische Philosophie* [Lecciones introductorias a la filosofía analítica del lenguaje] (1976) pueden reconocerse los contornos de este primer ensayo. Por entonces pensaba que la respuesta simple a la pregunta «¿Qué debo hacer?» —entendida como una petición de consejo, de acción racional— significaba en primer lugar «qué es lo mejor para ti», y a continuación «qué es *lo* mejor» (y esto significa obviamente «lo mejor para todos»). De esta forma parecía perfilarse un concepto de fundamentación práctica a dos niveles,

en el cual el punto de referencia del primer nivel era el propio bien, y el punto de referencia del segundo era la fundamentabilidad con respecto a todos.

Entre los años 1977 y 1979 trabajé en el Instituto Max Planck para la investigación de las condiciones de vida del mundo científico-técnico, en Starnberg, junto a Klaus Eder, Günter Frankenberg y Ulrich Rödel, en un proyecto de sociología jurídica sobre el desarrollo de las formas de fundamentación moral en la historia del derecho moderno. Cuando, en este contexto, me dediqué de nuevo a la problemática de fundamentación filosófico-moral, advertí las deficiencias de mi anterior intento de fundamentación práctica global y de la subsunción de la moral en un concepto de razón (véase p. 81); sin embargo, estaba todavía tan anclado en mi anterior concepción, que a partir de entonces adopté el siguiente punto de vista: cuando fundamentamos moralmente una acción y luego la norma correspondiente, lo que fundamentamos es la aceptación de que la acción —o la norma— es buena (es decir, buena para todos). En las extensas discusiones que mantuve por entonces con Jürgen Habermas quedó patente la contraposición de mi concepto de fundamentación semántica y el suyo, de carácter pragmático (comunicativo). Yo creía posible aprehender el sentido de una expresión moral en el sentido de la palabra «bueno», y tanto entonces como posteriormente me pareció que la pretensión de fundamentación contenida en esta palabra es la que constituye la pretensión de fundamentación de un juicio o una expresión moral.

La crítica a Habermas (véanse pp. 123 y ss. de esta edición), así como mi propia contraposición semántica, las presenté por vez primera en esbozo en mi conferencia leída en el congreso sobre «Argumentación y derecho», celebrado en septiembre de 1978 en Munich, en

el cual expuse nuestro proyecto de sociología jurídica, y luego, de forma más detallada, en otoño de 1979 en una conferencia sobre «Lenguaje y ética». La conferencia de Munich apareció publicada con el título «Zur Entwicklung von moralischen Begründungsstrukturen im modernen Recht» [Acerca del desarrollo de las estructuras de fundamentación moral en el derecho moderno] en el *Archiv für Rechts- und Sozialphilosophie*, n.º 14 (Wiesbaden, 1980), pp. 1-20. Parte de la conferencia sobre «Lenguaje y ética» la pronuncié de nuevo en noviembre de 1979 en la Universidad de Zurich. Este fragmento apareció en traducción francesa en la revista *Critique*, n.º 37 (1981), pp. 1.038-1.074. El original alemán no lo he publicado, pues posteriormente, con motivo de mi primera lección en Berlín, durante el semestre de verano de 1980, que tuvo de nuevo por objeto la ética, advertí claramente el carácter insostenible e ingenuo de esta concepción semántica. Esta concepción fracasa simplemente porque de la mera significación de una palabra, es decir, de un *a priori* comprendido de forma meramente analítica, no puede deducirse nada moralmente sustancial.

Tuve así que desechar la concepción de que, cuando fundamentamos una norma como norma moral, lo que fundamentamos es su bondad. Esto me llevó a entender de forma totalmente distinta la cuestión de qué significa en general fundamentar una norma como norma moral. De esta forma se perfiló mi nueva concepción, que he plasmado en las lecciones primera y tercera de las «Tres lecciones sobre problemas de la ética».

La segunda de estas «Tres lecciones» está dedicada a una cuestión concreta, la cuestión de si existen procesos de aprendizaje específicamente morales. Esta cuestión me la planteé en el contexto del citado pro-

yecto de Starnberg. Frente a mis otros dos colegas de Starnberg, que aceptaban ampliamente la concepción de las etapas del desarrollo moral de Piaget y Kohlberg, que se refiere a un desarrollo desde la perspectiva del observador, se me antojó la posibilidad de concebir pasos de aprendizaje más simples en primera persona sin una teoría de etapas, en la cual el criterio comparativo es exclusivamente el mejor fundamento (desde la perspectiva del propio sujeto). Debo a mis conversaciones con Ulrich Rödel sugerencias esenciales para la concepción aquí desarrollada. La versión original de esta segunda lección la presenté en una conferencia leída en España, en un congreso de ética celebrado en Madrid en otoño de 1979, y posteriormente en Lima; a comienzos de 1980 la ofrecí también en traducción al alemán con el título «Der Absolutheitsanspruch der Moral und die historische Erfahrung» [La pretensión de absoluteidad de la moral y la experiencia histórica] en sendas conferencias pronunciadas en las universidades de Hamburgo y Munich.

Las «Tres lecciones», en su forma actual, las pronuncié en marzo de 1981 en el marco del Seminario Christian Gauss de Princeton. Las he traducido al alemán para publicarlas aquí, sólo con algunas modificaciones menores. El original inglés no está publicado, pues ya pronto me asaltaron graves dudas con relación a aspectos esenciales de la concepción desarrollada en las lecciones primera y tercera. Pero sólo con motivo de las severas objeciones razonadas en la tesis de habilitación de Ursula Wolf, presentada en la Universidad libre de Berlín en la primavera de 1983, con el título *Über den Sinn moralischer Verpflinchtung*[1] [Acerca

1. Publicada en 1984 con el título *Das Problem des moralischen Sollens*, De Gruyter, Berlín.

del sentido del deber moral], no pude demorarlo ya más y conseguí la obligada distancia, que hizo posible una nueva y cabal revisión, como la que he realizado ahora en las «Retractaciones». Después de todas las transformaciones, la concepción que aquí he formulado en esbozo es la primera que al escribirla ya no tuve la sensación de que se trataba de un *tour de force*, y también la primera que merece ser publicada.

Las modificaciones con respecto a las «Tres lecciones» son tan decisivas, considero ahora erróneos aspectos tan esenciales de éstas, que su publicación en la forma que le he dado sólo es conveniente porque tanto la tesis de Ursula Wolf como el último ensayo de filosofía moral de Jürgen Habermas, «Ética del discurso. Notas sobre un programa de fundamentación»,[2] polemizan con ellas. No hubiera publicado las «Tres lecciones» sin las «Retractaciones». Por otra parte, las «Retractaciones» están escritas de tal modo que sólo se pueden comprender parcialmente si antes se han leído la primera y tercera de las «Tres lecciones». He intentado hacer de la necesidad virtud, y he presentado, en vez de pensamientos firmes, el proceso evolutivo de mis ideas y dudas.

Los otros dos trabajos de este libro ya estaban publicados, pero era difícil el acceso a ellos. El artículo sobre Rawls lo leí en un congreso sobre él —y en presencia de él— que tuvo lugar en abril de 1976 bajo la dirección de Gunther Pätzig en la Fundación Werner-Reimer en Bad Homburg. El original inglés apareció en *Analyse und Kritik*, n.º 1 (1979), pp. 77-89. Existe

2. Jürgen Habermas, «Diskursenethik — Notizen zu einem Begründungsprogramm», en J. H., *Moralbewusstsein und kommunikatives Handeln*, Frankfurt del Main, 1983, pp. 53-125. Versión española de R. García Cotarelo: *Conciencia moral y acción comunicativa*, Península, Barcelona, 1985, pp. 57-134.

un manuscrito de réplica de Rawls, pensado para la originalmente prevista publicación de las actas de este congreso. Con respecto a mi posición hacia Rawls, pueden verse también mis breves declaraciones a la revista *Zeit* del 4 de marzo de 1983 en la sección «100 libros especializados».

La conferencia «Antike und moderne Ethik» [Ética antigua y moderna] la pronuncié el 11 de febrero de 1980 en Heidelberg, en un coloquio celebrado con motivo del 80 aniversario de Hans-Georg Gadamer. Fue publicada en 1981 en *Die antike Philosophie in ihrer Bedeutung für die Gegenwart* [La filosofía antigua y su significación para la actualidad], editado por Reiner Wiehl en 1981 en Heidelberg (Actas de la Academia de Ciencias de Heidelberg, pp. 55-73), y se reproduce en este libro sin modificaciones.

Los trabajos de filosofía moral aquí publicados deben importantes sugerencias a las discusiones y conversaciones mantenidas en el Instituto Max Planck para la investigación de las condiciones de vida del mundo científico-técnico, fundado en 1970 y disuelto en 1982, al que pertenecí de 1975 a 1980, por lo que siento una especial deuda hacia Rainer Döbert, Jürgen Habermas, Gertrud Nunner-Winkler y Ulrich Rödel. A ellos y a todos los amigos de la época de Starnberg quiero dedicar el presente libro.

ERNST TUGENDHAT

Berlín, septiembre de 1983

OBSERVACIONES SOBRE ALGUNOS ASPECTOS METODOLÓGICOS DE *UNA TEORÍA DE LA JUSTICIA*[1] DE RAWLS

Si se compara la teoría de la justicia de Rawls con las otras dos teorías éticas contemporáneas tradicionales más importantes, la utilitarista por un lado y la kantiana por otro, se advierte un notable contraste con respecto al contenido y método de la teoría de Rawls. Desde el punto de vista del contenido, la teoría de Rawls va dirigida contra el utilitarismo, y a este respecto Rawls está —y así se considera él mismo— cerca de la concepción kantiana. Por el contrario, en sus convicciones metodológicas, Rawls se enfrenta a una concepción que se basa en «el análisis de los conceptos morales y el *a priori*», y opina que la tarea de una teoría moral consiste en ofrecer una teoría de nuestros «sentimientos morales» (p. 51). Una teoría moral debe ser contrastada en una clase de *hechos*, «nuestros juicios ponderados en equilibrio reflexivo» (p. 51). Rawls

1. John Rawls, *A Theory of Justice*, Cambridge (Mass.), 1972. (Hay traducción castellana de M. Dolores Gonzales, México, 1978. La paginación de las citas se refiere a la versión original.)

entiende que ésta es «la concepción de los autores clásicos, al menos hasta Sidgwick (p. 51). La tradición de autores clásicos a la que se une aquí Rawls es principalmente la tradición utilitarista. Es discutible su relación con Aristóteles, que sugiere en una nota a la citada afirmación. R. M. Hare, quien, si bien no es citado aquí, parece ser el verdadero objetivo de estos y otros pasajes similares, no es en modo alguno el primer filósofo que elabora su teoría ética a partir del análisis del significado de «bueno» y otros análisis conceptuales; lo mismo puede decirse de Kant y, si bien de modo muy diferente, de Aristóteles.

Una posible explicación del hecho de que Rawls esté metodológicamente tan cerca de su principal adversario —el utilitarismo— es que el propio Rawls pertenece ante todo a la tradición utilitaria. La disputa entre la teoría de la justicia y el utilitarismo parece ser una disputa entre parientes. Pero la teoría de Rawls está próxima al kantismo no sólo en cuanto al contenido; para justificar su teoría, Rawls utiliza un instrumento formal específico, una teoría contractual, y podemos preguntarnos si éste no es un elemento heterogéneo entre los demás supuestos metodológicos. Es cierto que, según Rawls, es precisamente este instrumento el que le permite «dejar a un lado las cuestiones de significado y definición y avanzar en el desarrollo de una teoría sustantiva de la justicia» (p. 579). Igualmente hay que conceder que une este elemento teórico-contractual de forma admirable con la concepción metodológica del equilibrio reflexivo. Los verdaderos principios de la justicia son según Rawls aquellos que serían elegidos en una posición, que es descrita como «posición original», pero sólo estaría justificado dar esta significación a la posición original si «los principios que se eligiesen coinciden con nuestras pondera-

das convicciones de justicia o las amplían de forma aceptable» (p. 19).

Sin embargo, la significación precisa y la fuerza justificativa de la posición original para las conclusiones de Rawls es precisamente lo que ha ocasionado mayores dificultades a los lectores críticos de *Una teoría de la justicia*. Por eso no parece demasiado atrevido suponer que el descuido de los análisis conceptuales, que deriva en Rawls de su concepción del equilibrio reflexivo, haya tenido un perjudicial efecto sobre la claridad de lo que se propone o consigue con el concepto de posición original. En el presente capítulo voy a abordar la cuestión en dos etapas. En primer lugar, voy a someter a examen crítico la concepción rawlsiana de la teoría moral y el concepto de equilibrio reflexivo. En la segunda parte trataré acerca de las repercusiones que tuvo el rechazo por parte de Rawls de las consideraciones conceptuales y analíticas sobre su concepción de la posición original como instrumento de justificación.

I

Según Rawls, la tarea de la filosofía moral consiste en hallar principios que respondan a nuestros «juicios morales ponderados». Esta formulación —añade— no constituye más que una primera aproximación, pues es probable que una persona modifique algunos de sus juicios morales ponderados a la luz de principios y, en particular, a la luz de principios diferentes. Tiene así lugar una recíproca acomodación de los juicios reflexivos y los principios, y cuando este proceso alcanza una detención provisional, Rawls habla de un «equilibrio reflexivo» (pp. 20; 48).

No resulta nada fácil entender esta concepción. Rawls explica que hay que considerar a «la teoría de la moral de forma similar a cualquier otra teoría» (p. 578). Esto parece presuponer que todas las teorías son básicamente similares. Rawls cita la lingüística (p. 47), la física (p. 49), la matemática (p. 51) y la teoría filosófica de la justificación de la inferencia deductiva e inductiva (p. 20). Sin embargo, la relación entre principios y hechos difiere en cada uno de estos casos de forma significativa con respecto a todos los demás. Una teoría lingüística tiene un objeto —el hablante competente— que ya se rige por principios o reglas, mientras que en el caso de una ciencia como la física, los datos en cuanto tales no tienen nada que ver con los principios; estos principios (leyes) sólo están en la teoría. Lo que parece proyectar Rawls para la teoría moral está, al menos a este respecto, más cerca de la lingüística que de la física. Pero incluso en el caso de la lingüística no sería muy correcto hablar de un equilibrio reflexivo en el sentido antes citado. El propio Rawls admite que «no podemos esperar una revisión esencial de nuestra noción de gramática correcta a la vista de una teoría lingüística» (p. 49). Pero esta diferencia entre la teoría lingüística y la teoría moral no parece ser tan contingente como la concibe el propio Rawls; quizá tenga que ver con el hecho de que la teoría moral, tal como la describe Rawls, se expresa en primera y segunda persona. Rawls llama a esto el «aspecto socrático» de la teoría moral (p. 49; p. 578). Resulta obvio que sólo si los datos atañen a la misma persona que hace la teoría pueden cambiar los datos a la luz de los principios que formula esta teoría.

Naturalmente, se puede construir una teoría de los sentimientos morales, muy similar a la descrita por Rawls, sólo con la diferencia de estar formulada en

tercera persona. En este caso, a la teoría resultante le faltaría simplemente el aspecto del equilibrio reflexivo, los datos no se modificarían con respecto a los principios, y estaríamos ante una simple teoría empírica. Cualquier teoría psicológica o antropológica del sentido de justicia de un grupo o sociedad sería de este tipo. La comparación con la teoría lingüística que ofrece Rawls parece adaptarse a este otro tipo de teoría moral, diferente al de Rawls.

Si esto es así, podríamos estar en condiciones de arrojar más luz sobre la concepción rawlsiana de una teoría moral preguntando por qué, en el caso de una teoría moral, la teoría en primera y en segunda persona puede tener una estructura significativamente diferente de la teoría en tercera persona, cosa que no sucede en el caso de la lingüística. ¿Por qué no tenemos una lingüística «socrática»? Y ¿por qué no tenemos un motivo especial para hacer lingüística en primera persona, pero podemos tener un motivo especial para formular una teoría moral en primera persona? Una primera respuesta parece obvia: nuestra competencia lingüística no mejora mediante la reflexión en sus principios, mientras que nuestro sentido de justicia puede mejorar por una reflexión semejante. Esto explica un aspecto de la concepción rawlsiana de la teoría moral que no he mencionado aún. Rawls dice: «una concepción de la justicia ... consiste ... en que todo encaje en una visión coherente» (p. 21). Esta teoría de la coherencia de la justificación moral constituye obviamente un corolario de la concepción del equilibrio reflexivo. No hablaríamos de una teoría de la coherencia, por ejemplo en lingüística ni en cualquier otra teoría empírica, porque en toda teoría semejante los principios tienen que concordar con los datos y no se plantea la posibilidad de un reajuste recíproco.

Sin embargo, todo esto sigue estando, todavía, demasiado en la superficie. Parece cierto que el hecho de que podamos tener una teoría significativamente diferente en primera persona en ética, pero no en lingüística, está ligado al hecho de que la reflexión puede mejorar nuestro sentido moral, pero no nuestra competencia lingüística, y también parece cierto que esta mejora tiene algo que ver con una mayor coherencia, pero subsiste la cuestión de conocer cuál es la razón de estas conexiones.

Sin embargo, esto es todo lo que pude hallar por medio de una mera dilucidación de la concepción de Rawls. Aunque tuve que introducir algunas distinciones que el propio Rawls no formula y aunque estas distinciones me parecen probar que la tesis de Rawls, de que puede considerarse una teoría moral igual que cualquier otra teoría, no es cierta ni siquiera para su propia concepción de la teoría moral, todo esto sólo pretendía ayudarnos a comprender la propia concepción de Rawls, y espero que hasta aquí él pueda estar de acuerdo conmigo. Pero si es así, parece difícil evitar un nuevo paso que no puede entenderse ya como una mera aclaración de la concepción de Rawls, sino que mostraría que, si se aclara debidamente, esta visión da pie a otra concepción diferente.

Vuelvo así ahora una vez más a la diferencia entre una teoría moral y una teoría lingüística. Pero todavía no hemos mencionado la más obvia diferencia entre el objeto de ambas teorías. Esta diferencia está implícita en el uso que hace Rawls de la expresión «*juicios* ponderados». Lo que Rawl denomina los «hechos» con los que tiene que tratar una teoría moral son una cierta especie de creencias, las creencias sobre lo correcto o justo. El carácter discursivo de estos hechos resulta oscurecido cuando se habla de *sentimientos* morales.

Ahora bien, las creencias o —por usar la expresión de Rawls— los juicios tienen la peculiaridad de que están vinculados a una pretensión de verdad o, si parece preferible, una pretensión de validez. La expresión lingüística estándar de una creencia o juicio es lo que se denomina un enunciado asertórico y, por definición, tales enunciados pueden ser verdaderos o falsos. Por supuesto, es controvertido si los juicios de valor o los juicios normativos pueden ser «verdaderamente» verdaderos o falsos. Pero no puede ser controvertido que son, por así decirlo, «fenomenológicamente» verdaderos o falsos. Por limitarme a la expresión favorita de Rawls, «justo», es obvio que los enunciados que expresan un juicio o creencia de que tal y tal cosa es justa o injusta tienen todas las características de cualquier otro discurso asertórico. Cuando expresamos lo que consideramos justo o injusto, utilizamos adverbios tales como «realmente», «verdaderamente», «aparentemente», «verosímilmente»; decimos cosas tales como: «Antes creía que esto era justo; después empecé a dudar de si realmente lo era y ahora ya sé que no lo es», etc.

Este hecho de que lo que Rawls denomina nuestro sentido de justicia consiste en un sistema de creencias contiene, creo yo, la explicación de por qué existe una significativa diferencia entre una teoría moral en primera y segunda persona y una teoría moral en tercera persona, pero esta diferencia resulta ser ahora mucho más profunda de lo que podría parecer según la formulación de Rawls. Lo importante es que lo que Rawls denomina los hechos de la teoría son, en este caso, hechos ligados a una pretensión de verdad. Para las personas de cuyos juicios se trata no son sólo hechos en los que se contrasta una teoría, sino que, al ser creencias, para ellas son datos susceptibles de ser contrastados. Un aspecto de cualquier sistema de

creencias es, obviamente, que tiene que ser coherente; si es autocontradictorio, no puede ser mantenido. Pero por sí solo esto no puede explicar la importancia que tiene aparentemente la reflexión sobre los principios para una teoría moral en primera persona. Tenemos que distinguir entre diversos tipos de sistemas de creencias. Las creencias sobre las cuestiones de hecho se justifican característicamente, de forma directa o indirecta, mediante la observación. Por otra parte, los juicios morales, en el caso que puedan ser justificados —y, al menos, pretenden serlo—, sólo pueden justificarse por principios. Así pues, la razón por la que los principios pasan a ser tan importantes en moral desde el punto de vista de las propias personas que formulan los juicios morales es que, al parecer, desempeñan un papel central en el proceso de justificación. Parece así que Rawls, si piensa en una teoría moral en primera persona, ha puesto el carro delante de los bueyes. No son los principios los que han de ser contrastados con los juicios morales particulares, sino al revés. Para no ser mal entendido, me apresuro a añadir que aun así queda abierta la posibilidad de una teoría moral en tercera persona, y ésta es, como cualquier otra teoría, una teoría en la que son los principios propuestos los que tienen que ser contrastados con los juicios particulares de las personas cuyo sentido de justicia está siendo estudiado. Pero obviamente una teoría semejante es una teoría, no de lo que es justo, sino de lo que las personas estudiadas creen que es justo.

Quizá fui demasiado dogmático al afirmar que los juicios morales sólo pueden justificarse mediante principios. Lo que quiero decir no es más que, si formulamos una teoría moral en primera persona, debemos advertir que nuestros juicios morales son expresiones que, por su propio sentido, no constituyen un tribunal

de apelación sino que están necesitados de él. La cuestión primordial para cualquiera que empiece a reflexionar sobre sus juicios morales es la cuestión de cómo puede justificarse este tipo de juicios. Mediante su no justificada asimilación de la teoría moral en primera persona a la teoría moral en tercera persona, Rawls se ha cegado a esta cuestión. Pero en este caso carecen de peso sus ataques a los análisis del significado. Si no existe problema de justificación, efectivamente no necesitamos ya un medio de abordarlo con seriedad. Pero si existe, podemos preguntarnos cómo se va a afrontar la cuestión del método de justificación de un tipo de enunciados sin un análisis del significado de estos enunciados.

Queda una esclarecedora referencia que hace Rawls en relación a su concepto de equilibrio reflexivo que no he mencionado aún. Se trata de la referencia a la teoría aparentemente similar de Nelson Goodman relativa a la inferencia deductiva e inductiva (p. 20, nota).[2] Según Goodman, «los principios de inferencia deductiva se justifican por su conformidad con la práctica deductiva aceptada» y «tanto las reglas como las inferencias particulares se justifican mediante su mutuo acuerdo» (Goodman, 1965, pp. 63, 64). Efectivamente esto suena bastante parecido a la concepción del equilibrio reflexivo de Rawls. La tesis de Goodman no es, sin embargo, incontrovertible, pero de cualquier modo hay una notable diferencia entre esta tesis y la de Rawls. Los «hechos», por utilizar la expresión de Rawls, consisten, en el caso de Goodman, no en juicios, sino en los procedimientos utilizados para justificar juicios. Transferir la idea de Goodman al caso de la teoría mo-

2. Nelson Goodman, *Fact, Fiction and Forecast*, Indianápolis, N. Y., 1965.

ral llevaría a un programa diferente al defendido por Rawls. El programa aspiraría ahora no a la justificación de principios morales, sino a la justificación de los métodos de justificación. Consistiría en el análisis de las reglas de la argumentación moral válida. No voy a determinar ahora si lo mejor que podemos hacer en este caso es, por analogía con lo que dice Goodman, justificar los principios de la argumentación moral válida contrastándolos con las prácticas aceptadas de argumentación moral; de cualquier modo, en el caso de la argumentación moral estaríamos mucho menos seguros de qué habríamos de considerar «prácticas aceptadas» que en el caso de la inferencia deductiva e inductiva. La alternativa a esta concepción sería una concepción como la de Hare: que las reglas de la argumentación moral válida derivan de la estructura lógica de estos enunciados. En el presente contexto voy a dejar abierta esta cuestión, porque mi disputa con Rawls no es porque no esté de acuerdo con la respuesta que da, sino porque ni siquiera plantea la cuestión y pone en su lugar una teoría de la coherencia. Por supuesto, es posible poner en duda que los juicios morales puedan justificarse de modo alguno; es posible mantener que su pretensión de verdad sólo les da apariencia de ser justificables y que no existen procedimientos de decisión para respaldar esta pretensión. Pero a su vez esta afirmación sólo podría fundarse en un análisis de la estructura lógica de estos enunciados. Rawls no adopta ni una actitud positiva ni negativa con respecto a esta cuestión: simplemente la pasa por alto.

Antes de concluir esta parte de mi artículo, permítaseme dar un paso hacia la reconciliación. Sería erróneo pensar que, a resultas de lo dicho, hay que abandonar la noción de equilibrio reflexivo. Lo único que

habría que hacer es interpretarla de diferente modo. Efectivamente, los juicios morales ponderados constituyen el punto de partida de cualquier reflexión moral, pero su valor es heurístico y no el de un tribunal de apelación.

Por ejemplo, Kant, en la primera sección de la *Grundlegung*, empezó por el análisis de nuestros «juicios morales ponderados»; a continuación, en la segunda sección, comprobó el resultado de este análisis mediante un análisis del concepto de acción incondicionalmente buena. La prueba procedió entonces en la dirección contraria a la defendida por Rawls. La segunda sección era para Kant la decisiva, y también para nosotros, que no compartimos ya todos los «juicios morales ponderados» de la época de Kant, es esta parte de su teoría moral la que ha seguido siendo valiosa. Lo que hace Kant en la segunda sección de la *Grundlegung* puede mostrar también la insostenibilidad de la afirmación de Rawls de que las cuestiones de significado y definición son inútiles para zanjar los problemas morales sustantivos. Éste es un argumento *ad hominem*, pues en el § 40 Rawls acepta las conclusiones sustantivas de Kant sin preocuparse por su derivación formal en Kant.

II

Podría parecer que la teoría contractual de Rawls nos ofrece en la práctica lo que parece negar en la teoría: un método de justificación. Pero, por supuesto, no hay aquí inconsistencia alguna, porque Rawls puede incorporar fácilmente su teoría contractual en su doctrina del equilibrio reflexivo. De hecho, mantiene que este procedimiento de justificación sólo puede justifi-

carse a su vez mostrando que su resultado encaja con nuestros juicios morales ponderados. Así, la teoría contractual de Rawls es en cierta medida neutral con respecto a la controversia que he abordado en la primera parte de este capítulo. Un filósofo que no esté de acuerdo con la afirmación de Rawls de que las reglas de razonamiento moral están justificadas si conducen a nuestros juicios morales ponderados podría estar aun de acuerdo con la teoría contractual de Rawls como marco general para la argumentación moral, que tendría que justificar independientemente.

Rawls llega incluso a satisfacer a medias a un filósofo así. Dice «que hay un considerable acuerdo en que los principios de justicia deberían haberse elegido en ciertas condiciones» (p. 18), y justifica su concepción de la posición original intentando mostrar no sólo que los principios elegidos en esta posición encajan con nuestros juicios morales ponderados, sino que esta posición también satisface aquellas condiciones que generalmente se consideran características del «punto de vista moral» (p. 120). Rawls parece otorgar a estas condiciones características del punto de vista moral un estatus similar que el que da a nuestros juicios morales ponderados. Aunque, según he podido ver, no es muy específico en relación con este punto, presumo que diría que no sólo tenemos juicios ponderados sobre cuestiones morales particulares sino también sobre las condiciones del razonamiento moral, y una teoría válida debería llegar hasta el equilibrio reflexivo por ambos lados.

Pero estas condiciones del razonamiento moral tienen obviamente un carácter abstracto y —*pace* Rawls— conceptual; pertenecen a la línea de clarificación que seguiría un filósofo que quisiera indagar si la descripción que hace Rawls de la situación inicial pue-

de considerarse una posición adecuada para la formulación de argumentos morales válidos. Por supuesto, un filósofo así tendría que seguir esta línea mucho más atrás de Rawls, quizás hasta el punto en que pudiera decidirse la cuestión por un análisis conceptual o lógico. Pero yo no voy a intentar hacerlo. No deseo seguir estas reflexiones sobre las condiciones de la argumentación moral más atrás del propio Rawls, o al menos no mucho más atrás, porque no pretendo realizar una crítica desde fuera.

Lo que quiero analizar en esta parte del capítulo es la cuestión de si la tendencial aversión a un enfoque analítico y conceptual que deriva de la concepción metodológica de Rawls no ha tenido efectos perjudiciales sobre la forma en que éste introduce la posición contractual. No pretendo decir que estos efectos o incluso la actitud negativa hacia las cuestiones conceptuales sea una consecuencia *necesaria* de la concepción metodológica de Rawls. Del hecho de que un argumento *a priori* sobre la validez del razonamiento moral tendría que haber sido conceptual, no se sigue que si no se intenta esta argumentación *a priori* no se necesiten los análisis conceptuales. Y para semejante concepción formal como la posición contractual de Rawls podría esperarse que el enfoque conceptual hubiese sido especialmente importante. Pero tampoco quiero exagerar ni insinuar, por ejemplo, que Rawls es conceptualmente poco claro. Lo que quiero mantener es que su introducción de la posición original no es suficientemente analítica para ser valorada adecuadamente.

Tiene un carácter considerablemente sintético por el hecho de ser un esquema multilateral, y Rawls no ha explicado paso a paso cuál de sus aspectos se sigue de las condiciones que supone son generalmente carac-

terísticas de la argumentación moral y qué aspectos ha introducido por otras razones; y ha hecho muy poco para mostrar la superioridad de su concepción comparativamente a otras concepciones que también satisficiesen aquellas condiciones de la argumentación moral. Al parecer, Rawls consideró suficiente indicar que *a*) muchos aspectos de la posición original concuerdan con aquellas condiciones, y *b*) que los principios elegidos en la posición original concuerdan con nuestros juicios morales ponderados. Por supuesto, esta situación debe resultar especialmente insatisfactoria para aquellos que quisiéramos considerar la propuesta de Rawls como una propuesta de la verdadera condición del razonamiento moral válido, pero además quiero probar que tiene también dudosos efectos sobre el punto de vista del propio Rawls relativo a la consonancia con nuestros juicios morales ponderados.

Una suposición fundamental de Rawls que no quiero discutir es que los principios de la justicia y los principios morales en general no son algo dado, en una presunta intuición por ejemplo, sino algo a lo cual llegamos activamente, en un acto de elección en determinadas condiciones. Estas condiciones de elección moral están circunscritas para Rawls en lo que él denomina «situación inicial» y por su ulterior especificación mediante la interpretación filosóficamente más favorable» denominada la «posición original» (p. 121; p. 146). Estrictamente hablando, en las condiciones de la posición original sólo se eligen los principios de estructura básica de la sociedad. Rawls concibe una «secuencia de cuatro fases» de creciente concreción de los problemas a decidir (§ 31).

Lo que resulta oscurecido por la forma en que Rawls introduce la posición original es que esta introducción representa en sí un acto de elección. La posi-

ción original tiene que ser *adoptada* como la *mejor* posición desde la cual decidir los principios morales en comparación con otras posibilidades como, por ejemplo, la teoría del observador imparcial (cf. pp. 184 y ss.). Cuando Rawls presenta las razones por las que deberíamos adoptar la posición original como la posición más adecuada en la que elegir los principios de justicia, está operando en una fase previa a la primera de sus cuatro fases. Por supuesto, esta fase cero de la elección moral no se caracteriza por un velo de ignorancia, pues el velo de ignorancia es una de las varias cosas que se eligen en esta fase. Tampoco se trata de una situación hipotética, pues, una vez más, la condición hipotética de la situación inicial es algo que, en la fase cero, es un objeto y no una condición de la elección. Por último, el tipo de elección precisa en una fase cero no es una «elección racional» en el «sentido restringido estándar de la teoría económica», característico de la elección a adoptar en la situación inicial (p. 14). Como hay que considerar a la deliberación necesaria en la fase cero como el fundamento de la filosofía moral, la afirmación de Rawls de que hay que concebir la «filosofía moral» como «parte de la teoría de la elección racional» (p. 172) no concuerda con su posición, al menos en este primer paso fundamental.

Así pues, la elección necesaria en la fase cero debe someterse a ciertas condiciones. Pero éstas no pueden estar determinadas por determinadas condiciones subjetivas (como ignorancia, racionalidad, etc.), sino sólo por el tipo de cosa a elegir. La cosa a elegir parece ser ésta: una representación adecuada del punto de vista moral. Las condiciones de la elección de la fase cero son, por tanto, las características definitorias del punto de vista moral; pero éstas pueden determinarse de dos modos. Pueden obtenerse por análisis lógico de

lo que puede significar justificar las proposiciones morales y, como he dicho, no voy a seguir esta dirección, pues es contraria a la adoptada por Rawls; o bien se pueden simplemente escoger, como hace Rawls, aquellas condiciones que parezcan aceptarse generalmente como características del punto de vista moral.

Pero como Rawls no ha presentado explícitamente lo que yo denomino fase cero, no ha empezado por donde habríamos esperado que lo hiciese, a saber, por una enumeración completa de estas condiciones. De este modo ha dejado sin aclarar qué aspectos de la posición original derivan de estas condiciones y qué aspectos ha elegido por otras razones. Esta falta de claridad sólo puede superarse reuniendo otras cosas relevantes que dice Rawls en diversos lugares. Lo más cerca que llega Rawls a una enumeración de estas condiciones es la enumeración de las «limitaciones formales del concepto de correcto» en el § 23. Las más importantes de estas «limitaciones formales» son la «generalidad» y la «universalidad». Según entiende Rawls estos principios, no parecen implicar imparcialidad; y, sin embargo, esta idea la subraya Rawls en diversos lugares como una condición esencial del punto de vista moral (por ejemplo, en relación con la introducción de la posición original, pp. 12 y 18). Estoy menos seguro del peso que otorga a la condición de «autonomía» (los principios tienen que ser «autoimpuestos», p. 13). Es una de las virtudes de la posición original que satisface esta condición, y esta condición no parece ser satisfecha por la teoría del observador imparcial, a pesar de lo cual cuando Rawls la analiza (§ 30) no la critica por este motivo. Si no incluimos la condición de autonomía, el punto de vista moral podría caracterizarse sumariamente como el punto de vista en que se elegirían estos principios de acción que van en interés

de todos. Se incluiría la condición de autonomía, si reformulamos esto caracterizando el punto de vista moral como aquella condición de la elección según la cual sólo son elegibles aquellos principios en los que todos pudieran estar de acuerdo.

Estas caracterizaciones son extremadamente toscas y precisarían una elaboración más minuciosa. Lo importante es que no basta con caracterizar la fase cero como mera enumeración de varias condiciones; tenemos que definir el punto de vista moral mediante alguna caracterización global. En contraste con los diversos modelos hipotéticos, como el modelo contractual o el modelo del observador ideal, el punto de vista moral no representa una situación de elección hipotética, sino la situación de elección moral de nuestra vida real (cierto es que incluso esta opción contiene elementos hipotéticos, cuando digo, por ejemplo, que hay que elegir aquellos principios en los que *pudiera* estar de acuerdo todo el mundo, pero la propia elección no es hipotética). No debería ser controvertido que la filosofía moral no puede *empezar* por una situación hipotética, sino sólo con el punto de vista moral como fenómeno de nuestra vida real.

El segundo paso consistiría en mostrar que en esta fase cero tenemos razones para adoptar una posición hipotética que sirva como representativa del punto de vista moral. Esto lo ha omitido Rawls. Lo que ha mostrado fue simplemente que la posición original incorpora las *mismas* condiciones características del punto de vista moral. Pero no ha mostrado explícitamente por qué es *preferible* cambiar la situación de elección de la fase cero a la posición original. De este modo, no queda claro si las razones de este cambio son *a*) razones que mejoran la propia perspectiva moral o *b*) razones prácticas o *c*) razones que tienen que ver

con la materia especial de la elección de los principios para la estructura básica de la sociedad, pero no quizá para otras elecciones morales. Y, por supuesto, Rawls no ha intentado sopesar las ventajas de su propuesta frente a sus concebibles desventajas.

En lo que resta de este capítulo sólo voy a esbozar algunas respuestas a estas cuestiones. La diferencia más básica del modelo contractual desde el punto de vista moral original es que permite separar el acto de elección de la consideración de los intereses de todos; se consigue imparcialidad no por la intención de llegar a un acuerdo o por determinado esfuerzo intencional simultáneo al acto de elección, sino por la aplicación anterior de un velo de ignorancia, con el resultado de que la elección pueda ser ahora una «elección racional» que sólo aspire a la propia ventaja, y con el resultado ulterior de que hablar de un «acuerdo» es realmente redundante, pues el acuerdo sería «unánime» (p. 139: «Por ello, podemos considerar la elección en la posición original desde el punto de vista de una persona seleccionada al azar»).

Parece que el principal motivo por el que Rawls considera a la posición original preferible al punto de vista moral original es que permite concebir la teoría de la justicia como «parte de la teoría de la elección racional» (p. 16), que parece ser algo más manejable que la elección Racional con una «R» mayúscula de la que tendríamos que hablar en la fase cero. Sin embargo, estaría todavía por probar si, en la práctica, la elección racional, cuando se realiza en la posición original, conduce realmente a resultados que sean mejores de algún modo a los que se habrían conseguido a partir de la fase cero. La principal prueba es aquí obviamente la justificación que hace Rawls del «principio de la diferencia» mediante este método, y luego volveré a

este problema. En cualquier caso, la ventaja de ser capaz de aplicar la teoría de la elección racional sería una ventaja práctica; mejoraría no nuestro concepto de lo justo sino los procedimientos de decisión para llegar a resultados justos. Ésta sería, obviamente, una ventaja que no debemos subestimar.

Otro rasgo en favor de la posición original que Rawls menciona a menudo es que permite la aplicación de la idea de «pura justicia procedimental» [o procesal: *procedural* (T.)]. La pura justicia procedimental se da cuando no existe un criterio independiente para el resultado correcto: por el contrario, hay un procedimiento correcto o ecuánime tal que el resultado, sea cual fuere, es igualmente correcto» (p. 86). Según puedo ver, la pura justicia procedimental sólo es necesaria cuando no se dispone de procedimientos de decisión más directos. Por ello no resulta adecuada para una clarificación de nuestra noción general de justicia, aunque constituye un instrumento moral limitado pero importante, adecuado para la decisión de ciertos problemas políticos y morales pero no de otros. Si la aplicabilidad de esta noción fuese una prerrogativa de la posición original, tendríamos aquí otra ventaja práctica decisiva y además una ventaja limitada a determinadas materias. La razón por la que la posición original parece especialmente apta para procurar una pura justicia procedimental es que este tipo de justicia supone un acuerdo previo en seguir ciertas reglas. Pero no hay razón por la que estos acuerdos, ya sean hipotéticos o reales, no pueden ser alcanzados directamente y *ad hoc* desde el punto de vista moral original. La concepción de la situación inicial basa toda la moralidad en un hipotético acuerdo preliminar. El punto de vista moral original no concibe de este modo el concepto de 'correcto' en general, sino que lo deja

abierto para determinarlo de este modo en aquellos casos en que es preciso.

Vuelvo ahora al problema del «velo de ignorancia». Al parecer, Rawls ha mezclado aquí varios aspectos diferentes. Introduce el velo de ignorancia en su habitual forma sintética y de golpe sin explicar por qué razón son necesarias las diversas partes de este velo. La única justificación que da respecto a la totalidad del velo (pp. 12 y 136) realmente sólo puede servir de justificación a una parte de éste: para asegurar la imparcialidad hubiera bastado que en la situación inicial todos ignorasen su propia identidad.[3] Rawls supone, además, que en la situación inicial todos deben ignorar las «circunstancias particulares de su propia sociedad» (p. 137). Una razón de esta suposición es que las «cuestiones relativas a la justicia social surgen tanto entre generaciones como en el seno de éstas» (p. 137), pero esto sólo no hubiese sido un motivo suficiente para exigir que en la situación inicial incluso «nos esté clausurado el curso de la historia» (p. 200). Estas nuevas limitaciones no derivan de la exigencia de imparcialidad sino porque «sin estas limitaciones de conocimiento, el problema de regateo [*bargaining*] de la posición original resultaría extraordinariamente complicado» (p. 140). Tenemos aquí otro rasgo especial de la posición original que no corresponde al punto de vista moral como tal sino que se añade por razones prácticas. Y una vez más, estas limitaciones parecen apropiadas sólo para ciertos problemas morales, aunque puedan ser los más fundamentales: en el modelo de

3. Véase R[ichard] M[ervyn] Hare, «Rawls' Theory of Justice», en *Philosophical Quarterly*, 23 (1973), pp. 144-155, 241-251; también en [y cit. según] Norman Daniels (ed.), *Reading Rawls. Critical studies on Rawls' «A Theory of Justice»*, Oxford, 1975, pp. 81-107, aquí pp. 89 y ss.

cuatro fases esta parte del modelo se elimina gradualmente (§ 31).

Sorprendentemente dice Rawls que en la última fase —«la aplicación de reglas a casos particulares»— el velo de ignorancia se levanta por completo (p. 199). Esto debe de ser un error si el tipo de elección en esta fase final ha de ser todavía de carácter autointeresado y si, no obstante, el resultado ha de ser imparcial. En la cuarta fase sólo puede levantarse por completo aquella parte del velo que había sido añadida por razones prácticas.

De nuevo podemos preguntarnos si el adicional velo de ignorancia que no deriva de la exigencia de imparcialidad constituye realmente una ventaja de la posición original. Como equivale simplemente a una decisión de no tener en cuenta todos aquellos hechos que parecen irrelevantes para la solución de un problema, esto puede conseguirse por supuesto directamente también desde el punto de vista moral, y probablemente mejor, porque en el punto de vista moral no estaríamos desligados de una vez por todas de toda la información; la cuestión de *qué* hechos son irrelevantes para la elección, por ejemplo de los principios básicos de justicia, podría quedar abierta durante el proceso de deliberación. En Rawls parece haber, por una parte, una tendencia a ignorar desde el principio todos los aspectos de la vida social cuyo valor comparativo no es cuantificable. Aquí parece ser el método de elección racional el responsable del velo de ignorancia.

Veamos ahora si todos estos rasgos adicionales que distinguen a la posición original del punto de vista moral constituyen realmente ventajas o no: *si* son ventajas, son ventajas prácticas, no morales; son introducidos para hacer más manejables las decisiones. Esto era de esperar, porque sería contradictorio pensar que

los rasgos estrictamente morales del punto de vista moral pudiesen mejorarse cambiando sus condiciones de elección, pues aquello que entendemos por «moral» o «correcto» se define por el punto de vista moral. La idea de Rawls de que podríamos «definir el concepto de correcto diciendo que algo es correcto si y sólo si satisface los principios que se habrían elegido en la posición original para designar cosas de este tipo» (p. 184 y p. 111) es una *petitio principii* y oscurece el hecho de que la adecuación moral de la posición original debe valorarse desde el punto de vista de la fase cero. El concepto de correcto sólo puede definirse diciendo que algo es correcto si y sólo si es el resultado de un proceso de decisión que empieza en la fase cero (que por supuesto no he definido adecuadamente). Lo mejor que podría decirse para la posición original es que el modelo de cuatro fases de Rawls es el procedimiento de decisión más operativo para los problemas morales que se plantean en la fase cero. Y en este caso estaría justificada la concepción de la justicia como equidad.

Pero ahora hay que abordar la cuestión de si las ventajas prácticas que posiblemente se consiguen por el cambio de la fase cero a la posición original no se pagan en una pérdida de sustancia moral. No resulta evidente de por sí el que si separamos la idea de una decisión moral (imparcial) en los dos componentes de una decisión autointeresada más la ignorancia de la propia identidad, se llega al mismo resultado. La concepción contractual, con su insistencia en un (aunque no sea más que hipotético) acuerdo *inicial* introduce en los problemas de la justicia un elemento de retraso temporal que no está contenido en el punto de vista moral original y en nuestra concepción ordinaria de justicia. Es este retraso temporal el que permite a

Rawls aplicar la teoría de la elección racional, pero los problemas de probabilidad que entran aquí con su específica contrapartida psicológica —la expectativa de las oportunidades y la disposición a [asumir] los riesgos— no parecen tener nada correspondiente en un juicio moral normal, excepto cuando por la naturaleza del caso no podemos sino adoptar los métodos de la pura justicia procedimental. Hay un largo camino desde la opinión de que hay que recurrir a una pura justicia procedimental, cuando no podamos hacer nada mejor, a la afirmación de que todo el problema de la justicia social es un problema de la justicia como equidad.

Rawls pensó que la trasposición del problema de la justicia en un problema de elección racional le daba el arma decisiva contra el utilitarismo. Pero varios críticos han indicado que parece ser un error que lo más racional a hacer en la situación inicial sea aplicar el principio maximin y optar así por una sociedad igualitaria.[4] Si estos críticos están en lo cierto, lo que se seguiría de las premisas de Rawls sería la concepción utilitaria. ¿Probaría esto que el utilitarismo es correcto y el igualitarismo incorrecto? Seguramente no, pues el punto de vista moral favorece claramente el igualitarismo. Por lo tanto, lo que parece seguirse es, más bien, que la posición original no constituye un modelo adecuado del punto de vista moral. Supongamos que alguien dice: «En la posición original, yo optaría por un sistema social regido por el principio de utilidad, porque esto maximizaría mis oportunidades; pero mo-

4. D. Lyons «Nature and soundness of the contract and coherence arguments» en *Reading Rawls*, pp. 141-167, aquí pp. 165 y ss.; véase también Hare, *ibidem*, pp. 102 y ss.; B. Barber, «Justifying justice: problems of psychology, politics and measurement in Rawls», en *Reading Rawls*, pp. 292-318, aquí pp. 296-298.

ralmente rechazo un sistema semejante por ser injusto». Según Rawls, sería autocontradictorio decir algo así, pero no parece serlo e incluso puede ser cierto.

Creo que la trasposición de la situación de elección moral original en una situación de elección autointeresada tiene otras consecuencias que muestran que se pierde algo de la sustancia moral. Una de éstas se refiere a la defensa de la libertad de conciencia. En la posición original difícilmente se puede defender tan directamente como hace Rawls la importancia de este derecho (p. 206). ¿Por qué personas sólo interesadas en sí mismas tendrían que apreciar algo como una conciencia moral? Por otra parte, si partimos del punto de vista moral original como el que he caracterizado, empezamos por considerar a todo individuo como persona moral, por tanto como sujeto y no sólo como objeto de deliberación moral.

Los dos últimos argumentos eran, en parte, argumentos *ad hominem*: mostrarían que la posición original conduce a unos resultados que no concuerdan siquiera con los «juicios morales ponderados» de Rawls. Sin embargo, la intención tanto de estos argumentos como de los anteriores no era descartar la posición original, sino apelar a una evaluación analítica de esta concepción.

(1976)

ÉTICA ANTIGUA Y MODERNA
(en el 80 aniversario de Gadamer)

Los aniversarios no son más que fechas ocasionales; constituyen la oportunidad para desear felicidad a una persona apreciada o querida cuando todavía es joven; cuando ya es de edad avanzada, son la ocasión para agasajarla como εὐδαίμων (feliz). De tratarse de filósofos quienes en semejante jornada se reúnen, éstos tienden obviamente a hablar *sobre* lo que los demás sencillamente hacen, y esto quisiera hacer yo hoy, hablar sobre una antigua pero abandonada cuestión: qué es la felicidad o εὐδαιμονία.

Esta pregunta, formulada en el aniversario de un filósofo, le da a uno —me da a mí— oportunidad de despojarse de la máscara objetivista de las disertaciones de los filósofos, como haría cualquier persona con otra en la discusión de un tema cualquiera. Quisiera también aprovechar la oportunidad de hablar explícitamente en primera persona, formular la pregunta del mismo modo que se me plantea en la perspectiva ocasional en la que me encuentro, y abordarla también en segunda persona, por consiguiente en conversación con Hans-Georg Gadamer.

Toda conversación filososófica, en tanto en cuanto tiene por objeto mejorar la propia comprensión del

asunto, en tanto en cuanto intenta comprender lo que el otro tiene que decirle a uno, es hermenéutica en el sentido de Gadamer. Una conversación filosófica, que por consiguiente ya siempre es hermenéutica en cuanto tal, tiene que comprender además su tema, si quiere ser precisamente una conversación con Gadamer, hermenéuticamente, es decir, tiene que comprenderlo históricamente. Con anterioridad no he tenido por costumbre abordar hermenéuticamente las cuestiones filosóficas, pero creo haber llegado a un punto en el que los límites de mi anterior y ahistórico modo de proceder muestran —y lo reconozco— que necesito el diálogo con la hermenéutica, lo que por su parte me remite al diálogo con la historia. La cuestión que aquí quiero poner a discusión, que atañe a la necesidad de una confrontación de la ética moderna con la antigua, es *un* ejemplo en el cual esto me ha quedado muy claro. Pero antes de entrar en esta temática hermenéutica, resulta obligatorio dilucidar lo que en filosofía puede significar, en general, la hermenéutica. O bien, para formularlo de nuevo desde la perspectiva de la segunda persona: si se quiere discutir con Gadamer sobre una cuestión filosófica, no sólo se tiene que contemplar esta cuestión en la dimensión hermenéutica, sino que se tiene que discutir también con él sobre el sentido [*Sinn*] de la hermenéutica, y por consiguiente sobre qué es, en términos generales, qué se hace o se debería hacer, cuando se aborda una cuestión filosófica desde la perspectiva hermenéutica.

En esta cuestión previa sólo puedo entrar de forma sumaria, presentando de manera meramente dogmática, cuál es, en contraste con la de Gadamer, mi propia concepción. Podría preguntarse si una tal posición dogmática no es precisamente lo contrario de la discusión a la que aspiro. Pero aquí defiendo, por lo demás por

vez primera, un μακρὸς λόγος (discurso libre), y en éste se puede mantener abierto lo dicho para la discusión ulterior, por lo que sólo se llama dogmático (pues se distingue como mera opinión subjetiva), a la vez que el aseguramiento [*Absicherung*] argumentativo e incluso la escenificación de un diálogo ficticio en una conferencia pueden servir precisamente para el aislamiento con respecto el diálogo real.

Tanto antes como ahora,[1] lo que no me convence de la concepción de Gadamer de la hermenéutica filosófica es que éste la concibe según el modelo del encuentro con la obra de arte. Esto significa en primer lugar que, para Gadamer, la situación de partida hermenéutica es el encuentro con el otro, encuentro que, obviamente, sólo puede ser productivo cuando en él se reflexiona al mismo tiempo sobre uno mismo; por el contrario, yo quiero tomar como situación de partida la reflexión sobre sí, sobre la propia comprensión, de donde resulta la necesidad de la profundización histórica. Esta primera distinción podría parecer un mero cambio de énfasis, pero con ella va relacionada una segunda distinción, que atañe al concepto de verdad. La verdad de diferentes obras de arte —en el supuesto de que pueda hablarse aquí de verdad— no se relativiza recíprocamente. Por el contrario, se fundamenta en el sentido mismo de la verdad enunciativa [*Aussagenwahrheit*] el que diferentes concepciones filosóficas de un asunto se relativizan mutuamente y, en mi opinión, sólo se puede concebir adecuadamente una hermenéutica filosófica cuando consigue comparar esta relativización con la verdad enunciativa, y esto no significa dar un sentido relativista. Una relativización sin más, en la que la ver-

1. Véase mi recensión en *The Times Literary Supplement*, 77 (1978), p. 565. Allí se explica con más detalle lo que expongo a continuación.

dad se fragmenta en un firmamento de posibilidades —de modo que siempre se puede decir, se puede contemplar tal como dice A, y contemplar también como dice B, etc., va en contra de la propia pretensión de verdad de A, B, etc., y me parece que la concepción de Gadamer conduce tendencialmente a un relativismo de este tipo. Gadamer está lejos de otra forma de relativismo, el relativismo que rechaza precisamente la propia pretensión de verdad. Ésta es la relativización histórica en el mismo sentido de cuando se dice que, en las condiciones causales x, y, z parece verdad lo que dice A, y en las condiciones u, v, w, lo que dice B. Esta segunda concepción incluso no sería ya hermenéutica sino explicativa. A resultas de ella, ya no se tomaría en serio la pretensión de verdad de A, B, etc. —hecho en el que siempre insiste Gadamer—, no se hablaría ya *con* A, B, etc., sino *sobre* él en tercera persona. Pero lo que me parece significativo de este relativismo explicativo y no hermenéutico es que se puede integrar en la perspectiva de la primera persona. En efecto, tan pronto como el sujeto a que nos referimos cobra conciencia de que su concepción anterior del objeto dependía de determinadas condiciones, ya no sigue siendo, tras esta constatación, alguien como el observador en tercera persona, sino que, cuando de nuevo se orienta al conocimiento del objeto, no puede hacer otra cosa que formular una nueva concepción de éste, libre de esta relatividad. De este modo se da la posibilidad de una progresiva desrelativización. Sin embargo, esto presupone que se atenga a la pretensión de verdad absoluta. Sólo desde la perspectiva de una pretensión de verdad absoluta conduce la conciencia de la determinación de una concepción, en vez de a un relativismo, a su desrelativización.

Ahora puedo ya explicar cómo entiendo yo aproxi-

madamente el modo de proceder hermenéutico en filosofía. El punto de partida (por así decirlo, la tesis) es siempre la referencia objetiva [*Sachbezug*] pre-hermenéutica en primera persona, derivada de una pretensión de verdad absoluta. Lo que se concibe desde esta perspectiva se pone en cuestión tan pronto como advierto la determinación de mi concepción, o bien cuando encuentro otra concepción del mismo objeto. En cierta medida, ésta es la antítesis. Y la síntesis consiste en una nueva concepción unitaria de la cosa, concepción en la cual está integrada la otra perspectiva; esto admite los dos casos límite, que se crea poder conocer que la otra concepción, o bien la propia, es sencillamente falsa o que en ambos casos una supere a la otra; lo más normal es que se tenga que desarrollar una tercera concepción, frente a la cual tanto la propia concepción original como la que se ha encontrado en una experiencia hermenéutica resultan unilaterales. La concepción que aquí presento podría parecer una hegelianización de la hermenéutica. Pero, tal como yo lo concibo, el proceso hermenéutico avanza paso a paso, ninguna experiencia hermenéutica se puede anticipar de antemano, y aparte de esto el punto de partida es en todos los casos la propia situación contingente. En donde yo quisiera insistir frente a Gadamer es en que la experiencia hermenéutica se construye a partir de la ingenua experiencia pre-hermenéutica. Así pues, sólo en tanto en cuanto se conserve la pretensión absoluta de la experiencia ingenua se puede, en mi opinión, evitar el relativismo.

Con esta observación preliminar sólo quería poner en claro la perspectiva desde la cual quiero abordar un problema hermenéutico como el de la ética antigua y moderna. De lo dicho se desprende que el título correcto hubiera sido «Ética moderna y antigua». El punto de

partida es la ingenua instalación en el seno de la propia tradición moderna; yo no puedo, como el historiador, situarme comparativamente en cierto modo por encima de ambas tradiciones, pues para el que filosofa no existe semejante tercer lugar. Estando como estoy en la tradición moderna, salgo al encuentro de otra, la tradición antigua. Ésta no es simplemente una tradición extraña, sino una tradición de la cual la modernidad se ha separado conscientemente. Por consiguiente, tendré que examinar en primer lugar el fundamento de este paso. Si es convincente, de ello se sigue que, contenga lo que contenga la tradición antigua que vaya más allá de la moderna, en lo que ésta se muestre como unilateral, no podemos volver simplemente a la tradición antigua, sino que podemos valorar aquello que encontramos en la antigua, sólo como un indicio de un problema que ha seguido abierto entre nosotros; y no podemos abordarlo de nuevo más que sobre nuestro terreno, con nuestros métodos de fundamentación. Obviamente, hay opiniones accesorias fuertemente subjetivas que, como todo lo que sigue, no voy a fundamentar suficientemente. Lo que pretendo con lo que aquí voy a decir no es que lo consideren verdadero, sino al menos significativo y, por consiguiente, discutible.

En primer lugar está obviamente la cuestión de si con el título global de «ética antigua» y «moderna» se puede delimitar en general algo identificable. Voy a dejar a un lado ahora esta cuestión por cuanto en cada uno de los dos ámbitos, la ética antigua y la moderna, fijo una cuña —en Aristóteles y Kant— y a continuación digo: y con ello me refiero también al ámbito adyacente; con ello no determino exactamente la extensión del campo adyacente. Naturalmente, una decisión semejante, la de tomar a Kant como punto de partida, sólo es posible cuando mi propia autocomprensión

ingenua, de la cual tengo que partir en definitiva, está en la tradición kantiana. Para justificar la importante implicación de que puede considerarse a la ética kantiana como representativa de la ética moderna, sólo quisiera decir lo siguiente. Me parece que la única otra tradición ética moderna independiente de la kantiana es la del utilitarismo; en aquel lugar de la confrontación con la ética antigua que en definitiva me interesa, van sin embargo Kant y el utilitarismo en el mismo barco. Por el contrario, las demás tradiciones éticas que se han desarrollado en la modernidad, como la hegeliana o la ética material de los valores, son por su parte reacciones a la ética kantiana, que por lo demás han vuelto a abandonar un plano de fundamentación alcanzado por Kant y, en mi opinión, irrenunciable; sobre ello volveré más adelante. Aparte de esto, estas tradiciones éticas, en su reacción a Kant, han recurrido a una determinada interpretación de la antigüedad, con lo que su inclusión en el contexto actual hubiera producido excesivas complicaciones. La otra decisión, la de orientarme paradigmáticamente a Aristóteles para referirme a la ética antigua, resulta ciertamente menos problemática y carece de consecuencias para la siguiente argumentación, por lo que si no están de acuerdo con ella pueden sustituir simplemente la expresión «ética antigua» por «ética aristotélica».

Siendo así que contemplamos las cuestiones éticas desde la perspectiva de Kant, ¿en qué medida puede el encuentro con Aristóteles llamarnos la atención acerca de una deficiencia en nuestra propia posición? En este lugar quisiera situarme una vez más en diálogo directo con Gadamer. Pues también en él encontramos una confrontación de la ética aristotélica con la kantiana, en particular en su conferencia de Walderberg del año 1961, que lleva por título «Über die Möglich-

keit einer Philosophischen Ethik» [Sobre la posibidad de una ética filosófica].[2] Tanto aquí como en otros lugares de su obra ocupa un lugar destacado lo peculiar del concepto aristotélico de φρόνησις (prudencia).[3] Mientras que en la «ética legal» de Kant parece deducirse lo correcto en cada actuación [*Handlungssituation*] del principio moral general, Aristóteles tiene un concepto del saber moral a resultas del cual, lo correcto en cada caso debe ser conocido de modo inmediato en la situación particular de entre lo que se encuentra en la concepción adecuada, es decir, inspirada en el principio. De acuerdo con su concepción de la hermenéutica, Gadamer contrapone globalmente la posición aristotélica a la kantiana. Si las concepciones filosóficas son como las obras de arte, en la práctica tiene que parecer ilícito realizar trasplantes. Pero cuando se trata de la verdad se tendría que poder convencer al propio kantiano de la mayor conveniencia de la concepción aristotélica en este punto, pero esto significaría o llevarle a abandonar su posición o mostrar que el problema de aplicación aristotélico se plantea también precisamente cuando se parte de un principio no partidista, como es el imperativo categórico. También se podría intentar probar que lo que en cada caso es no partidista no se puede deducir sencillamente del principio, sino que exige una facultad de discernimiento [*Urteilsvermögen*] no deductiva, ligada a la situación. A continuación sólo querría indicar la dirección en la que intentaría integrar en la posición kantiana la idea

2. Hans-Georg Gadamer, «Über die Möglichkeit einer Philosophischen Ethik», 1961, en *Kleine Schriften*, vol. 1, Tubinga, 1967, pp. 179-191.

3. Véase también: Hans-Georg Gadamer, *Wahrheit un Methode*, 3.ª ed. aumentada, Tubinga, 1972, pp. 295 y ss. (Hay trad. cast.: *Verdad y método*, Sígueme, Salamanca, 1977.)

de φρόνησις, tan convincentemente destacada por Gadamer.

Ciertamente, en la comparación de Kant con Aristóteles que encontramos en la conferencia de Walderberg, un segundo punto va unido a éste. Como es sabido, la capacidad de formular juicios morales correctos concretos —de φρόνησις— según Aristóteles, no es una facultad intelectual libre [*freischwebendes*], sino que depende de la disposición afectiva adecuada del ser humano que, por su parte, remite a una correcta educación. Gadamer describe esta vinculación como una vinculación «entre la subjetividad del saber y la sustancialidad del ser»,[4] y habla asimismo de la «fundamental sustancialidad de derecho y costumbre».[5] Si se ahonda ahora un poco en esta bastante clara vinculación con la crítica de Hegel a Kant, acaso podría formularse así la valoración de la diferencia entre Kant y Aristóteles en este punto: Kant insistió erróneamente sólo en la cuestión de la fundamentabilidad de lo correcto en el ámbito práctico; aquí, lo correcto sólo se puede reconocer si simultáneamente se pospone a la existencia previamente dada [*Vorgegebenheit*] de derecho y costumbre; y esto es precisamente lo que ha pretendido Aristóteles. No creo reproducir aquí con exactitud la concepción de Gadamer, y quizá no hago más que construirme un espantajo, pero lo necesito en este lugar para poder destacar de forma más clara lo que quiero decir. Y de nuevo me voy a contentar con una contraposición dogmática.

Lo que quiero decir es lo siguiente. 1. De la dependencia de la facultad del juicio moral con respecto a la disposición afectiva correcta, es decir, con res-

4. Hans-Georg Gadamer, *Kleine Schriften*, vol. 1, p. 187.
5. *Ibid.*, p. 186.

pecto a la educación correcta, no se sigue que, por su parte, la educación correcta no sea fundamentable; por consiguiente, no se sigue que deba posponerse la moral [*Moral*] a una moralidad [*Sittlichkeit*] previamente dada, es decir una moralidad cuya fundamentabilidad no sea ya más interrogable. 2. Encuentro inaceptable que se admita algo como correcto o bueno porque está ya dado de antemano en la costumbre, sin poderlo probar como correcto o bueno; ello iría no sólo en contra de una idea moderna de la filosofía, sino incluso de lo que ya desde Sócrates significa filosofía: un radical dar cuenta y razón [*Rechenschaft*]. 3. Puede decirse que el criterio de lo moralmente correcto que da Aristóteles, el sopesamiento [*Ausgewogenheit*], es tan indeterminado, que (sólo) encuentro su determinación fácticamente mediante lo dado de antemano en la costumbre del momento; pero creo que hay que distinguir dónde termina de hecho la posición de Aristóteles, en cierta medida a sus espaldas, y en qué consiste ésta según su propia intención. El propio Aristóteles nunca ha señalado como normativo lo dado de antemano. Ha formulado una pretensión de fundamentación, pero una pretensión que finalmente no pudo concebir. 4. La modernidad se distingue de la antigüedad en la radicalización de los criterios de demostración, y ello tanto en los juicios teóricos como prácticos. Para Kant no podía quedar sin determinar cómo hay que fundamentar una clase de juicios. Llega así a la distinción entre una fundamentación apriorística en sentido estricto y justificado y una fundamentación empírica, y también de una fundamentación metafísica, injustificada, de la que pueden darse criterios objetivos de validez general. Esto en particular puede requerir una revisión. La radicalización de la idea de fundamentación constituye no obstante un progreso en el sentido de la autonomía

y el dar-razón pretendidos desde los tiempos de Sócrates.

Por consiguiente, quiero separar los dos puntos que parecen confundirse en la comparación de Gadamer entre Aristóteles y Kant. Con respecto al segundo punto me distancio así de Gadamer ante todo en la valoración. No podemos estar de vuelta de la moderna pretensión de fundamentación. Obviamente, veo una vinculación entre el segundo y el primer punto, que yo describiría sin embargo de otro modo. Una proposición moral, en la que no se expresa lo que queremos sino lo que debemos hacer, no puede ser una proposición empírica, por lo que sería entonces una proposición teórica, y no de deber, con lo cual tendríamos lo que en inglés se conoce como *naturalistic fallacy*. Por consiguiente, sólo puede basarse o en una decisión, y no ser así fundamentable en general, o bien debe ser fundamentable *a priori*, y tratarse así naturalmente sólo de un *a priori* formal, no metafísico. Fue por tanto la radicalizada pretensión de fundamentación la que condujo a Kant a su principio moral apriorístico-formal. La forma más sencilla de desarrollar este proyecto es que valga *a priori* no sólo el principio mismo, sino que también las normas concretas generales a partir de él forman un corpus apriorístico, que entonces vale de una vez para siempre, lo que significa precisamente esto: ya no hay que abandonar la decisión de qué es lo correcto en cada caso concreto a una facultad de juicio situacional.

Lo que en su primer punto constata Gadamer como debilidad de Kant en comparación con Aristóteles es también, según mi descripción del segundo punto, una consecuencia, si bien no obligatoria, de este segundo punto. Pero ahora se materializa este —en mi opinión— punto fuerte de Kant en comparación con Aris-

tóteles y resulta imprescindible. Surge entonces la tarea de reexaminar la comprensión de Aristóteles de la «situacionalidad» [*Situationsbezogenheit*] del juicio moral, de nuevo explicada por Gadamer, *en* el plano de la radicalizada pretensión de la fundamentación kantiana, para integrarla luego en la concepción kantiana. Si, según creo debe formularse la cuestión, los juicios morales concretos no son simplemente derivables del principio moral, como Kant pensaba; si, por su parte, éstos no se pueden fundamentar *a priori*, ¿qué sucede entonces?, ¿dónde se fundamentan otramente que en el principio? Parece haber aquí dos posibilidades no excluyentes. La primera es que el juicio moral concreto en general no puede fundamentarse concluyentemente, que contiene un resto decisionista. El sujeto del juicio moral aspira ciertamente a la imparcialidad, pero no puede dar mayor fundamento al resultado concreto de su juicio situacional. O bien, podría indicarse que (y cómo) en la fundamentación del juicio situacional entran, además del principio *a priori*, también conocimientos empíricos, debiendo explicarse por consiguiente la estructura de esta experiencia y fundamentación específicamente morales. Podría pensarse que la concesión de un componente decisionista contradice la radical pretensión de fundamentación. Pero sería erróneo. La radicalizada pretensión de fundamentación exige meramente no dejar en la indeterminación el estatus de fundamentación; por ello, la concepción de que un juicio en general no es fundamentable o sólo lo es parcialmente, es un caso límite legítimo. El decisionismo está tanto en el terreno de la autonomía como el racionalismo.

Ahora querría sólo indicar la dirección en la que creo uno tendría que intentar apropiarse, como filósofo moral moderno que ha pasado por Kant, la idea de

φρόνησις de Aristóteles. Con ello sólo querría también preparar la perspectiva metódica, en la que ahora querría entrar en relación con otro aspecto de la filosofía moral aristotélica y antigua en general, singularmente adecuado para llamarnos la atención acerca de otro déficit de la ética moderna. Se refiere a la cuestión de cuál es en general el objeto de la ética. Es ya sabido que bajo la misma palabra «bueno» se esconden, en la ética antigua y en la moderna, temas de diferente tipo. La antigua pregunta por el ἀγαθόν, por el *bonum*, se refería a aquello que en cada caso es bueno para el individuo, su bien, aquello con lo cual su querer puede satisfacer plenamente la εὐδαιμονία. Por el contrario, el objeto de la ética kantiana o utilitaria se refiere a las normas intersubjetivas. La formulación de la pregunta de la ética antigua era: qué es aquello que quiero verdaderamente para mí; la de la moderna es: qué es aquello que debo hacer con respecto al otro. Esto no significa naturalmente que los griegos no conociesen esta temática de la ética moderna: tenían sin embargo para ella otra denominación: no τὸ ἀγαθόν (lo bueno), sino τὸ καλόν (lo bello). Esta temática moral, la de lo καλόν, también se plantea por supuesto en la ética antigua, pero de tal manera que está integrada en la pregunta por la verdadera felicidad. Es bien sabido cómo sucedió esto. La orientación impulsora [*Stossrichtung*] de la ilustración sofística consistió en desenmascarar como heterónoma la moral tradicional, oponiendo a ella la praxis autónoma como la realmente orientada al propio interés. Platón ha asumido esta crítica en la medida en que, igualmente, para él lo tradicionalmente dado no puede tener una pretensión normativa fundada, y también para él sólo puede comprenderse una praxis autónoma como aquella orientada hacia el interés propio bien entendido; la tesis conocida en

toda la ética antigua, que siempre fue normativa en ella, fue que precisamente lo καλόν es lo verdaderamente ἀγαθόν, y que por consiguiente sólo cuando nos comportamos moralmente satisfacemos nuestro propío interés bien entendido, tesis para la cual era obviamente esencial una firme distinción entre el interés propio aparente y el bien entendido, entre la felicidad aparente y la verdadera. Pero no sólo la problemática moral sino también el hecho notorio de que no estamos seguros de lo que realmente queremos y podemos pedir consejo a otros, ponían de manifiesto que hay que distinguir —como dice Aristóteles— entre una φαινόμενον βουλητόν (satisfacción aparente) y una βουλητὸν ἀληθές (satisfacción verdadera), entre aquello en lo que la voluntad se satisface propiamente y lo que la satisface genuinamente, y esto parece nuevamente exigir una medida independiente del querer fáctico. Se estaba muy cerca aquí de la referencia a una salud psíquica, e incluso los conceptos de τέλος (fin) y ἔργον (obra), y por consiguiente —quizá podríamos traducirlo así— de la idea de un perfecto desarrollo de sí mismo; y la tesis era por lo tanto: sólo quien lo ha alcanzado, sólo el psíquicamente sano, puede ser verdaderamente feliz.

En su concepción antigua, la ética se relaciona por consiguiente primeramente con la cuestión acerca de la verdadera felicidad y sólo secundariamente aborda la problemática moral. Por el contrario, si nos orientamos hacia Kant y hacia el utilitarismo, la ética moderna no sólo se refiere primariamente a la moral, por lo tanto a lo que los griegos llamaban τὸ καλὸν καὶ τὸ δίκαιον (lo bello y lo justo), sino que ya no incluye en general aquella pregunta primaria para los griegos, la pregunta por el ἀγαθόν, por la verdadera felicidad. ¿Por qué no? En Kant podemos encontrar claramente las razones. La pregunta por la real o verdadera feli-

cidad puede abordarse más por el lado subjetivo, en cuanto se distingue un determinado tipo de estado sensible como la verdadera felicidad; pero también puede abordarse de forma más objetiva, en tanto en cuanto se indica un criterio que no constituye por sí un estado sensible, sino una determinada concepción de la persona, por lo tanto quizás el ya citado perfecto autodesarrollo. Kant rechazó ambas posibilidades, la subjetiva y la objetiva. En primer lugar, Kant desecha cualquier recurso a un concepto de perfección en ética, porque semejantes conceptos son «vacíos e indeterminados» y con ellos uno no hace más que moverse «circularmente»:[6] con ello sólo se introduce en la imagen teórica del ser humano lo que ya se presupone práctico-normativamente. O bien, para resaltar más claramente la vinculación de la tesis kantiana con mi problema: no se puede dictar a la voluntad desde cualquier concepción descriptiva del verdadero ser humano, sólo en la cual éste pudiera satisfacerse verdaderamente. El lugar en que la voluntad se satisface verdaderamente sólo puede mostrarse en esta misma satisfacción, en el correspondiente estado sensible. Como no hay ningún tránsito legítimo del ser al deber ser, tampoco hay ningún tránsito legítimo del ser al querer. O, por decirlo de otro modo: acerca de la verdadera felicidad sólo puede decidir la propia felicidad.

Nos vemos remitidos por tanto a la variante subjetiva, pero ésta fue igualmente rechazada por Kant. Kant y Bentham convienen en que no se puede distinguir los placeres superiores o verdaderos de los

6. Kant, *Grundlegung zur Metaphysik der Sitten*, en *Gesammelte Schriften*, ed. de la Real Academia Prusiana de Ciencias [en lo que sigue citada como «Edición de la Academia»], vol. 4, p. 443.

inferiores. El argumento decisivo de Kant es el siguiente: mientras que en el ámbito de la moralidad hay reglas objetivas de validez universal acerca de lo que *debemos* hacer, no hay reglas de conducta objetivas y de validez universal para la consecución de la felicidad. La cuestión de qué acción es absoluta, es decir, moralmente buena, puede fundarse objetivamente; por el contrario, no puede fundarse objetivamente, con validez universal, la cuestión de qué acción es buena para mí, qué acción favorece mi bienestar. No puede fundarse ningún concepto de felicidad de un contenido determinado, y esto resulta evidente por cuanto o nos vemos remitidos al hecho subjetivo en el que se satisface nuestro querer y desear —y esto no es ninguna fundamentación objetiva—, o bien tenemos que recurrir a una medida independiente del estado subjetivo, con lo que nos exponemos a la crítica kantiana al concepto de perfección. Resulta así que la razón por la que en la ética moderna la cuestión acerca de lo intrínsecamente querido no sólo se pospone a la del deber, sino que en general caduca, es de nuevo la acusada pretensión de fundamentación. Además, obviamente no puede perderse de vista la orientación política que fundamenta esta concepción: la convicción básica a la concepción jurídica liberal de que debe dejarse a cada cual configurar su vida a su manera. Éste es un enunciado de deber que se deduce precisamente del imperativo categórico. Por supuesto, esta prohibición de intervenir en la autonomía del individuo no presupone que no existan principios objetivos fundamentales para la configuración de la vida, pero cuando se cree en estos principios se está a un paso de una dictadura moral.

¿Significa esto que podemos renunciar hoy a la que para los antiguos era la cuestión fundamental? Creo

que no. No voy a desarrollar aquí el argumento fundamental de la cuestión, a saber: que, en calidad de seres dotados de la capacidad de reflexión, no podemos dejar de plantearnos la cuestión relativa a nuestro verdadero bien, lo cual nos sitúa siempre en la tensión entre intereses fácticos y verdaderos, entre la verdadera y la supuesta felicidad. En su lugar, sólo querría indicar que la pregunta por el bien moral es, por su mismo sentido, una pregunta de orden superior, que por sí misma remite a la pregunta por lo que es bueno *para* cada cual, y esto en dos sentidos.

En primer lugar, en relación con el contenido. Pues es moralmente bueno aquello que va en interés no partidario de todos (también el imperativo categórico kantiano tiene finalmente este sentido). Sin embargo, esto implica que nosotros, para conocer lo moralmente prescrito, debemos saber qué va en interés de cada cual, y aquí no se trata en definitiva del interés fáctico, sino del interés bien entendido de cada cual. Si bien éste no puede ser anticipado —pues, de lo contrario, se pondría en entredicho la autonomía de los sujetos—, subsiste el interés moral en que los sujetos intenten conocer qué es lo que realmente quieren. Más claro aun que en el ámbito de la moral individual se aprecia esto en el ámbito de la moral social, el ámbito de la justicia. La justicia social significa igual reparto, pero igual reparto, ¿de qué? Creo que desde Kant hemos constatado que no basta la igualdad jurídica formal y, además, hemos experimentado que tampoco basta la igualdad económica o la igualdad de oportunidades, porque presupone una definición económica del bien individual. Esta definición mínima del bien parecía abandonar la configuración de la vida a la autonomía individual. Pero en realidad esta concepción de la justicia implica una determinada decisión previa relativa a la configu-

ración de la vida individual; además, la configuración de la vida individual se muestra cada vez más dependiente y ligada a las decisiones sociales globales, que por su parte se presentan con la pretensión de justicia. Pero esto significa que esta justicia no puede ya comprenderse como una justicia de oportunidades económicas, sino más generalmente como una justicia de oportunidades de vida o felicidad, lo cual presupone que en su concepción la reflexión debe entrar a determinar qué sea la felicidad. Naturalmente hablo de forma intencionada de «felicidad» y no empleo expresiones como «la vida digna de un ser humano», porque sobre ellas caería en seguida la crítica kantiana al concepto de perfección. Por difícil que sea formular la pregunta por la verdadera felicidad en el plano de la pretensión de fundamentación alcanzada con Kant, ya no podemos hoy considerar como Kant esta pregunta tan irrelevante como la pregunta por el bien moral; más bien, desde entonces ha mostrado precisamente nuestra experiencia relativa a la repercusión que sobre la concepción de la justicia social ha tenido el concepto moral kantiano —en mi opinión, de fundamentos irrenunciables—, que la concreción de este concepto moral no sólo tiene que referirse a las situaciones en general, sino que nos remite a la problemática de la felicidad.

Pero ésta parece imponerse ahora en un segundo sentido, a saber, en lo referente a la *motivación* para la acción moral y, por consiguiente, a la pregunta: ¿por qué querer ser moral? Los antiguos creían poder responder a esta pregunta sólo mediante el recurso a la verdadera felicidad, y me parece que en principio no es posible otra respuesta. Como es sabido, Kant ha intentado concebir la racionalidad [*Vernünftigkeit*] de lo moralmente prescrito como motivo posible de

la acción, para lo cual tuvo que pensar que esta racionalidad remitía a un estrato suprasensorial de la personalidad, dividiéndola así en dos estratos. Quien no esté en condiciones de seguirle en esto tendrá que, o abandonar la idea de una genuina acción moral, o bien esperar una respuesta a la pregunta por la verdadera felicidad que incluya la idea de moralidad [*Moralität*]. Si no se comprende erróneamente el concepto de felicidad, esto resulta incluso forzoso. Sería una equivocación contemplar la búsqueda de la felicidad como egoísmo o incluso como algo amoral desde el principio. Incluso en Kant se encuentra una definición según la cual «felicidad es el estado de un ser racional ... al cual, en el conjunto de su existencia, le va todo según su deseo y voluntad».[7] Pero la voluntad no se dirige necesariamente al bienestar de la propia persona, sino que por el contrario podría decirse más bien: obtengo bienestar cuando, aquello a lo que se dirige mi voluntad, alcanza satisfacción; y cualquiera que sea la fórmula de moralidad que se elija —por ejemplo, la kantiana, no considerar a nadie sólo como medio sino siempre al mismo tiempo como fin— significa que esto, por consiguiente acaso el bien del otro, cuando es algo que yo quiero por sí mismo, es un componente de mi concepción de la felicidad. Esta proposición [*Satz*] parece concluyente, pero naturalmente se trata de una proposición de la forma 'Si... entonces': *si* alguien obra autónomamente en sentido moral, entonces lo hace sólo porque quiere esto mismo, lo que significa que lo hace porque pertenece a su felicidad. Por el contrario, la cuestión es, primero, cómo debe estar

7. Kant, *Kritik der praktischen Vernunft*, en Edición de la Academia, vol. 5, p. 124. (Hay trad. cast.: *Crítica de la razón práctica*, por Emilio Miñana y Manuel García Morente, Madrid, 1975, p. 175.)

constituida nuestra concepción de la felicidad si tiene que incluir [*einschliessen*] a la moralidad y, segundo, si *sólo* es verdadera semejante concepción de la felicidad que incluya a la moralidad, si, por lo tanto, sólo somos realmente felices (una idea de la que Kant se burla) cuando —como creían los antiguos— nos comportamos moralmente.

Es por consiguiente la propia moral —y quiero decir la concepción moral moderna, que pasa por Kant— la que tanto por lo que respecta al contenido de lo moralmente prescrito como también con relación a la motivación moral, la que constituye en una necesidad filosófica la reformulación [*Wiederaufnahme*] de la antigua pregunta por lo verdaderamente querido, los intereses bien entendidos, la verdadera felicidad. Pero con ello quiero decir entonces: al igual que la problemática de la θρόνησις, sólo podemos retomar hoy la problemática de la εὐδαιμονία si lo hacemos en la moderna conciencia del problema, es decir, sólo de forma tal que con ello no retrocedamos con respecto a la pretensión de fundamentación después formulada. Esto no significa un neocartesianismo, como si sólo pudiésemos hablar de lo que se puede fundamentar con seguridad, y tuviésemos que callar del resto. No se trata de la seguridad de la fundamentación, sino de la claridad acerca del tipo de posibilidad de fundamentación.

Por consiguiente, ¿cómo se puede reformular —y ésta es la pregunta que quería formularles— la problemática de la felicidad en el plano de la conciencia metódica actual? Con ello tenemos que dejar a un lado aquel aspecto que, desde el punto de vista práctico, tiene una considerable relevancia y en particular no es en modo alguno trivial, pero en principio resulta completamente aproblemático: la cuestión de qué caminos son mejores para realizar una determinada concepción

de la felicidad. Mi problema no se refiere a los posibles errores en cuanto a los medios, sino en las metas de nuestro querer. Atañe a la cuestión de qué puede significar equivocarse con respecto a los propios intereses o necesidades auténticas.

La distinción entre necesidades o intereses verdaderos y falsos no es en modo alguno desconocida sino, por el contrario, muy extendida, pero hasta ahora no he encontrado ningún autor (lo cual puede deberse, por supuesto, a mi falta de lecturas) que se ocupe de esta distinción con plena consciencia de las dificultades metódicas y ofrezca un criterio de la verdad de los intereses que pueda afrontar la crítica kantiana. Está particularmente extendida en la crítica social la referencia a falsas necesidades, como por ejemplo hace Marcuse en *One dimensional Man* (1964). Pero por plausible —e, incluso, indiscutible— que sea el que aquello que consideramos nuestros intereses depende de las estructuras sociales vigentes, me parece que un acceso sociológico o filosófico-social no puede ofrecer ningún criterio de los intereses verdaderos; pues, ¿cómo se va a fundamentar la idea de que sólo en una determinada sociedad —por ejemplo, participativa y no jerárquica— se pueden desarrollar los verdaderos intereses? Ostensiblemente toda tesis semejante parte de una determinada concepción del hombre, sólo puede fundarse —como en Aristóteles— en una determinada concepción de la constitución correcta del individuo, por supuesto concebido como ser esencialmente social. Del plano sociológico nos vemos remitidos, por tanto, al plano psicológico. Pero entonces se plantea también aquí la cuestión de con qué derecho podemos destacar una determinada concepción psíquica como correcta o natural. La cuestión es si existe un criterio según el cual a una persona le van las cosas bien o mal, inde-

pendientemente de su bienestar fáctico presente y futuro, y que no esté afectado por la crítica kantiana al concepto de perfección.

Indudablemente hay *una* dimensión del ser humano para la cual disponemos de un criterio objetivo semejante, la salud y la enfermedad humanas. Para la afirmación «él está bien» (o «le va bien») disponemos por una parte de criterios subjetivos conformes con la definición kantiana de felicidad antes citada: está bien cuando su situación corresponde a su deseo y voluntad; pero con igual derecho empleamos también la afirmación tal y como la podemos emplear con relación a las plantas: está bien cuando está sano. Este criterio objetivo no puede reducirse al criterio subjetivo, pues alguien puede estar enfermo sin saberlo. Pero me parece que vale la relación inversa: sería irracional no querer estar sano y también lo sería querer sufrir dolor sin motivo alguno. En realidad, no estoy totalmente satisfecho con esta afirmación; habría que llenar aquí un vacío. Pero parece obvio que en la salud tenemos un criterio del bien objetivo, independiente de la voluntad y efectivamente normativo para ésta desde su propia perspectiva. Por eso tiene fundamento que los griegos, en la interrogación acerca de la felicidad, se hayan orientado hacia la salud y también hacia el concepto de función, hacia el ἔργον, pues la enfermedad significa menoscabo de la capacidad de funcionamiento.

La cuestión decisiva es entonces si —y de qué manera— podemos seguir aun a los griegos, haciendo también extensivo este concepto de salud o capacidad de funcionamiento a la constitución anímica. ¿Es la referencia a la salud y claridad anímica sólo una metáfora, o bien tiene un sentido más diáfano? Esta pregunta se ha vuelto particularmente apremiante en la época actual, pues hoy existe una terapia que precisa de un con-

cepto de salud anímica, el psicoanálisis; por supuesto, en la literatura psicoanalítica sobre esta cuestión se ha cobrado conciencia progresivamente de que toda fijación sustancial [*inhaltliche*] de lo que hemos de entender por salud anímica, podría depender de una determinada idea normativa, de una determinada concepción humana, lo que nos introduciría en seguida en aquel círculo al que apuntó Kant en su crítica del concepto de perfección.

Necesitamos por consiguiente un concepto formal de salud psíquica. Puedo explicar lo que quiero decir haciendo referencia una vez más a la concepción griega. Para la medicina griega, la salud consistía en armonía, equilibrio, mezcla correcta de sustancias corporales opuestas, y así era natural que Platón y Aristóteles concibieran también la salud anímica como equilibrio. ¿Se puede hacer también hoy? Se ha intentado. Así desarrolla Erich Fromm, en su libro *Man for himself* (1947) en consciente vinculación con la tradición aristotélica sobre una base psicoanalítica, una concepción a resultas de la cual la salud psíquica consiste —abreviadamente— en el equilibrio entre la propia independencia y la necesidad de los demás. La moderna orientación hacia la relación sujeto-objeto, es decir, hacia la intersubjetividad, permite a Fromm (como ya permitió también a Hegel y Kierkegaard) dar un sentido más determinado que Aristóteles al concepto de equilibrio o síntesis de opuestos, que puede parecer fácilmente una fórmula vacía. Pero por plausible que encuentre la concepción de Fromm, también está, tal y como está formulada, expuesta a la crítica kantiana.[8]

8. Aquí y en lo que viene a continuación estoy en deuda con Gertrud Nunner-Winkler. A ella le agradezco también la sugerencia del ensayo de Kubie.

Fromm ha descuidado aquí, igual que los griegos, una distinción conceptual. Si se dice que la salud —bien se trate de la salud corporal o la anímica— consiste en un equilibrio, no queda claro si esto debe ser una definición de la salud o bien una proposición empírica, a saber, que la salud se *fomenta* mediante un equilibrio constituido de tal y tal manera. En mi opinión, está claro que sólo puede ser esto último, y que la proposición tiene por tanto el estatus de una hipótesis causal. Pero esto presupone que disponemos de un criterio de salud y enfermedad independiente de esta determinación. También se da con respecto a la enfermedad física: la enfermedad es el menoscabo de la capacidad de funcionamiento. Resulta así patente qué es lo que falta tanto en los griegos como en Fromm: necesitamos un concepto formal de salud psíquica, con el que sea tan evidente como con la salud física decir que sería irracional no querer estar sano; semejante concepto no estaría ya expuesto a la crítica kantiana. Aunque el equilibrio puede determinarse siempre con una precisión cada vez mayor, no resulta en modo alguno evidente que sea un objeto necesario de nuestra voluntad (como tampoco cualquier otra concepción sustantiva), pero sería lógicamente un objeto necesario de nuestro querer si fuese empíricamente correcto que es una condición causal necesaria de la salud psíquica.

Por consiguiente, la pregunta es ¿hay un concepto de menoscabo de la capacidad de funcionamiento psíquico que se corresponda con el concepto de enfermedad física? Creo que en la literatura psicoanalítica puede encontrarse un concepto semejante, cierto que la mayoría de las veces asociado a otros contenidos, pero está elaborado rigurosamente en un ensayo de Lawrence Kubie, quien precisamente se propone esta tarea, hallar un criterio unívoco y general de compor-

tamiento enfermo. Según Kubie, la característica general de la conducta patológica es que se trata de una conducta compulsiva [*zwanghaft*], que se repite automáticamente «sin consonancia con la situación, el provecho o las consecuencias de la acción», por contraposición al comportamiento normal o sano, que es flexible y controlado a voluntad.[9] Naturalmente, esta conducta compulsiva también es querida en cierto sentido, pero es característico de ella que la persona tiene que actuar así, quiera o no. Semejante obrar no es accesible a la reflexión, no puede ser influido por ésta y por lo tanto no es libre, no elegido por mí mismo, no autónomo. Este concepto de autonomía es obviamente más formal y débil que el kantiano: se refiere no a una autonomía de la razón, sino de la persona en su conjunto, lo que significa la autonomía del querer por completo determinada sensiblemente, que sólo es racional en el sentido de la capacidad de reflexión. Con ello se habría conseguido un concepto formal de enfermedad psíquica que no sólo se refiere a un menoscabo de la capacidad de funcionamiento psíquico en general, sino más precisamente a un menoscabo de la capacidad de ejercicio del querer. De este modo se habría conseguido lo que se buscaba, un punto de vista independiente de las metas en cada caso subjetivas de nuestro querer y sin embargo normativo desde la misma perspectiva del querer. En tanto en cuanto seres volitivos [*als Wollende*] —en el sentido de libres electores [*frei Wählenden*]— siempre queremos no estar limitados en nues-

9. Lawrence S. Kubie, «The Fundamental Nature of the Distinction between Normality and Neurosis», en *The Psychoanalytic Quarterly*, 23 (1954), pp. 167-204; reimpreso en L. S. K., *Symbol and Neurosis. Selected Papers*, ed. por Herbert J. Schlesinger, Nueva York, 1978. Véanse aquí, especialmente, pp. 142 y 161.

tra libre elección. Normalmente no conocemos la medida de la 'compulsividad' [*Zwangshaftigkeit*] de nuestro querer; pero en la medida en que la conociésemos, sería irracional quererla —igual que es irracional querer estar enfermo—, y la vinculación es aquí aun más íntima, pues querer estar enfermo no es deseable sólo en sí; podríamos tener igualmente buenos motivos para querer estar enfermos como medio para otros fines, pero nunca para no querer elegir libremente.[10]

Con relación a la pregunta por la verdadera felicidad, se sigue pues que (como ya nos ha enseñado la filosofía existencial) sólo quiero algo de verdad cuando lo quiero realmente en el sentido de que lo elijo libremente. La referencia a verdaderas necesidades o intereses bien entendidos resulta ser así equívoca, pues da la sensación [aparente] de que en algún lugar existieran en sí estos verdaderos objetos de nuestro querer, y sólo tuviéramos que descubrirlos; la cuestión de lo verdaderamente querido se refiere así no a las metas de nuestro querer, sino al 'cómo' [*Wie*] del querer. La palabra «verdadero» es aquí un adverbio, no un adjetivo. Por supuesto habría que dilucidar empíricamente si no es así que, cuando se quiere de modo no forzado, la voluntad así concebida apenas puede ya orientarse a determinadas metas, y sólo por esta vía podría fundarse un punto de vista como el de Fromm, así como la referencia a las verdaderas necesidades.

No podemos delegar la pregunta por la verdadera felicidad —sustraernos por tanto nosotros, es decir,

10. Ello llevaría también a pensar el concepto unitario de enfermedad humana, no a partir de lo físico, sino de lo psíquico. También la enfermedad es una limitación, si bien periférica, del libre poder disponer de uno mismo. El poder disponer de sí mismo sería entonces el concepto fundamental, y no la capacidad de funcionamiento.

sustraer a nuestro querer— en una instancia objetiva, aunque esta instancia estuviese en nosotros. Quiero decir por tanto que la antigua pregunta por la verdadera felicidad no se ha vuelto obsoleta hoy, sino que sólo podemos encontrarle una respuesta formal, formal en un sentido afín a la respuesta que ha encontrado la interrogación por la moral en Kant. Con respecto ahora a la moral de concepción kantiana tendría que mostrarse en mi opinión qué y cómo la formalidad del principio no excluye la dependencia de la experiencia y la 'situacionalidad' [*Situationsbezogenheit*] de las normas fundadas en este principio;[11] así, la respuesta formal a la pregunta por la verdadera felicidad me parece que sólo representa el necesario punto de partida, y habría que preguntarse ahora, en primer lugar, cuáles son las condiciones psíquicas y las metas sustanciales posibles de semejante querer y, en segundo lugar, cuáles son las condiciones sociales que posibilitan estas condiciones psíquicas. En esta labor, sería interesante examinar qué representaciones sustantivas [*inhaltliche Vorstellungen*] de la respuesta aristotélica a la pregunta por la vida buena habría que reivindicar a partir de este planteamiento formal.

(1980)

11. Véase sobre el particular mi conferencia «Kann man aus der Erfahrung moralisch lernen?» (¿Se puede aprender moralmente a partir de la experiencia?) (pp. 99-122 de la presente edición).

TRES LECCIONES SOBRE PROBLEMAS DE LA ÉTICA

OBSERVACIÓN PRELIMINAR

El título de estas lecciones podría dar la impresión de que yo quisiera hablar sobre diferentes problemas de ética relacionados de forma más o menos libre. Sin embargo, no es ésta mi intención. Las tres lecciones se refieren a una única pregunta: la pregunta por la posibilidad de fundamentar nuestros juicios morales. Hablo de problemas —y esto en plural— sólo porque a mi entender esta pregunta se divide en otras preguntas parciales, y porque mis reflexiones no pretenden constituir una teoría y tienen más bien un carácter totalmente provisional y problemático.

En todos los demás tipos de juicios —por ejemplo, los científicos o los estéticos— se puede considerar el problema de la fundamentación como una cuestión puramente académica. Sólo en relación con la moral constituye el problema de la fundamentación una necesidad de la vida concreta. Ésta es la razón por la que, también históricamente en la sofística y Sócrates, la moral marcó el lugar en el que la reflexión filosófica (entendida ésta como un λόγον διδόναι, un dar cuenta y razón) se originó en la vida concreta. La pregunta

por la fundamentación de los juicios morales era ineludible y vuelve a reaparecer inevitablemente en el momento histórico en que las convicciones morales pierden su fundamentación religioso-tradicional anterior, percibiéndose acto seguido como históricamente relativas.

No se puede caracterizar adecuadamente por qué en tales épocas se vuelve ineludible la pregunta por la fundamentación si se comprende la pregunta moral fundamental de la moral tal como se formula a menudo a partir de Kant, es decir, en primera persona del singular como la pregunta «¿Qué debo hacer?». Naturalmente, el individuo puede padecer bajo la incertidumbre de qué debe hacer, pero esto también puede sucederle en relación con la incertidumbre de, por ejemplo, sus convicciones religiosas o con respecto a la carencia de tales convicciones. En cambio, lo que bajo las condiciones de relatividad histórica es no sólo una posibilidad sino una necesidad, es la fundamentación de los propios juicios morales con respecto a los demás, por el simple hecho de que la moral consiste en la recíproca exigencia de determinadas acciones y omisiones. Cuando nuestras convicciones morales divergen, estamos ante el hecho de que exigimos a los demás que restrinjan su libertad de modo que no les parece evidente; por eso, cuando no se nos someten simplemente, nos vemos abocados forzosamente a la cuestión de fundamentar estas exigencias. Y en el caso de que no haya de antemano convicciones morales comunes, la alternativa a la fundamentación será la fuerza. En cambio, cuando en cuestiones relativas a la ciencia, la religión, el gusto o la dirección de la vida privada se insiste en convicciones comunes, esto quiere decir que se están convirtiendo estas cuestiones en cuestiones morales, lo cual se puede hacer [*kann man*] aunque no se tiene

que hacer [*muss es nicht*] (y, naturalmente, constituye una cuestión moral la de determinar si es lícito [*ob es darf*]). Es entonces este hecho, que la necesidad de fundamentar las convicciones morales es de carácter intersubjetivo —y esto porque la moral consiste en exigencias intersubjetivas—, el que nos lleva a la idea de que la fundamentación de los juicios morales bien pudiera tener por su parte una estructura intersubjetiva. Ahora bien, no está claro de antemano lo que puede significar en general que la fundamentación de una clase de juicios deba ser de carácter estructuralmente intersubjetivo; pero precisamente por esto voy a dedicar la tercera lección a esta pregunta.

Si la fundamentación de los juicios morales es esencialmente algo que acontece entre diversos sujetos (individuales o colectivos), un aspecto de este problema lo constituye la circunstancia de que los diferentes sujetos tienen diferentes convicciones morales en razón de las diferentes condiciones históricas de su vida —tomando aquí la expresión «condiciones históricas» en un sentido amplio que se refiere no sólo a la historia colectiva, sino también a la biografía de los individuos. Un diálogo de fundamentación moral entre diferentes sujetos, es por tanto, un diálogo sostenido desde diferentes posiciones históricas. La concepción habitual es que la convicción de la relatividad histórica de los juicios morales y la convicción de la posibilidad de fundamentación se contradicen. Frente a este punto de vista, en mi segunda lección quiero abordar la cuestión de si, y cómo, le puede corresponder a la reflexión sobre las condiciones históricas de las propias convicciones morales una función positiva en la fundamentación de los juicios morales. Quisiera ver hasta dónde se puede avanzar en esta cuestión poniendo entre paréntesis el problema de la comunicación, para volver

de nuevo a ella desde un nivel más alto en la tercera lección.

El propósito de la primera lección es la clarificación previa del concepto de moral (o de un concepto de la moral) y de la cuestión de qué puede entenderse en general por fundamentación de los juicios morales.

1. El acceso semántico a la moral

Delimitar un concepto significa delimitar la comprensión de determinadas expresiones lingüísticas. Por consiguiente, nuestra primera tarea es un análisis del significado de las proposiciones morales. Sin embargo, parece estar hoy día pasado de moda un semejante acceso semántico, como el que predominaba en la filosofía moral inglesa y norteamericana de los años cincuenta y sesenta. Según Rawls —por citar sólo al autor más destacado—, el acceso semántico nos ha apartado de la problemática sustancial de la filosofía moral. ¿Es que en realidad sólo se ha tratado de una moda? Naturalmente, esto sólo lo podemos decidir examinando los argumentos esgrimidos. Lo importante es llegar a aclarar básicamente qué habla en favor o en contra de un planteamiento semántico, qué valor tiene un tal planteamiento, y hasta dónde llega.

A primera vista, la idea de un acceso semántico a la filosofía moral tiene que parecer absurda a toda persona que considere el problema sin prejuicios. «No puede usted creer en serio —se me objetará— que mediante un análisis de cómo se *habla* sobre lo bueno y lo malo, se vaya a obtener un criterio sobre lo que *es* bueno y malo. ¿No sería también absurdo pretender que en las ciencias empíricas se puede decidir lo que es verdadero y falso mediante un mero análisis de

cómo hablamos?» «Pero, ¿cuál es en su opinión el criterio —respondería yo— según el cual hay que decidir cómo se debe actuar y cómo no?» Mi interlocutor podría dar entonces una de las respuestas conocidas: el criterio son los mandamientos de la ley de Dios, o el sentimiento moral del individuo o de la comunidad, o lo útil para la comunidad, o la razón práctica, etc.; o bien podría unir alguna de estas respuestas de una u otra forma, o rechazarlas todas y adoptar una posición relativista o escéptica. Pero como quiera que responda, podría yo formular a continuación la contrapregunta de cómo pretende mi interlocutor fundamentar su concepción.

En el caso de que quisiera afrontar la pregunta (y no veo cómo podría evitarlo), la conversación pasaría a un segundo plano. En el primer plano se trataba del criterio conforme al cual fundamentamos nuestras convicciones morales; en cambio, se trata ahora de la pregunta de cómo fundamentamos el criterio fundante.[1] Sería erróneo creer que con ello se ha repetido simplemente la pregunta por la fundamentación, de modo que pudiéramos temer caer en un regreso al infinito. Se trata más bien de dos diferentes planos. Esto se puede apreciar fácilmente mediante el caso —comparable a este respecto pero más transparente— de las proposiciones empíricas a las que se ha referido mi propio interlocutor. Para una proposición empírica vale decir que su criterio de fundamentación es, en última instancia, la experiencia, y cuando en el segundo plano preguntamos por la fundamentación de este

1. Puede verse una aplicación y desarrollo de esta idea relativa a los planos de fundamentación en el manual de Friedo Ricken, *Allgemeine Ethik*, Kohlhammer, Stuttgart, 1983. (*N. del t.*)

criterio resulta que se funda en el significado de una proposición empírica.

¿Qué se puede desprender de este ejemplo de las proposiciones empíricas en relación con nuestro problema? Creo que nos muestra que mi interlocutor tenía toda la razón al afirmar que no debemos buscar en el lenguaje sino en la experiencia el criterio de fundamentación de las proposiciones empíricas; pero esta afirmación, por su parte, sólo se puede fundamentar mediante un análisis semántico de las proposiciones correspondientes.

¿Es inevitable concebir este segundo plano como un análisis semántico? Entiendo que sí. En efecto, el objeto que hemos de afrontar en la reflexión filosófica —ya se refiera a la ciencia, a la moral, a la estética o a la creencia religiosa— consiste inicialmente en expresiones lingüísticas; el primer paso no puede consistir, pues, en otra cosa que en esclarecer qué es lo que entendemos por el tipo correspondiente de expresiones lingüísticas; y cuando se trata de la pregunta de cómo hay que fundamentar un determinado tipo de proposiciones, nos vemos asimismo remitidos a la estructura semántica de dichas proposiciones. El análisis filosófico siempre se ha concebido como un análisis conceptual. El modo de acceso semántico se encuentra dentro de esta tradición y se diferencia de la tradición anterior sólo en la concepción —que yo tengo por bien fundamentada— de que a los conceptos no se puede acceder mediante una intuición intelectual, sino sólo en el modo de empleo de nuestras expresiones lingüísticas.

Por lo tanto, no se trata de una moda sino de un método al que —hasta donde alcanzo a ver— no hay alternativas. Pero, ¿cómo comprender entonces que un filósofo moral contemporáneo del tamaño de Rawls

haya podido repudiar el método semántico? La única explicación posible —caso de ser correctas mis consideraciones— es que él ha permanecido en el primer nivel. Y éste es en efecto el caso: Rawls presupone en la construcción de una «posición original» que el criterio de la moral es la imparcialidad, criterio que a su vez no fundamenta. La razón por la que Rawls se ha permitido presuponer simplemente este principio es que no se ha ocupado en absoluto del problema de cómo podía fundamentar sus convicciones morales. La tarea que él se había propuesto consistía en hallar principios que concuerdan con las nuestras —y, en último término, como es natural, con sus propias convicciones morales «ponderadas». No hay que menospreciar una tal concepción de la teoría de la justicia; quizá de la ética de Aristóteles también podría afirmarse que sólo tenía esta intención reconstructiva. Pero en todo caso está claro que una teoría semejante apunta a otra cosa (y podría decirse que a algo no tan de principio), distinta a una concepción de la filosofía moral en el sentido de una investigación sobre cómo fundamentar los juicios morales. En consecuencia, la circunstancia de que un análisis semántico no sea importante para una teoría semejante no dice nada contra la necesidad de un análisis de este tipo para la pregunta por la posibilidad de fundamentación.

Ahora bien, ¿cómo debemos comenzar el análisis semántico de las palabras y proposiciones relevantes? En relación con esta pregunta hallamos la segunda razón por la cual el modo de acceso semántico ha caído en descrédito entre los filósofos morales recientes. No resulta tan claro como les pareció a los representantes clásicos de la filosofía moral del análisis del lenguaje, en especial a Richard Hare, que haya una o dos expresiones estándar que aplicamos en los juicios morales.

Hare fundamentó su análisis de los juicios morales en un análisis general del significado de la palabra «bueno». Otros filósofos, en primer lugar Von Wright en su libro *The Varieties of Goodness* (1965), insistieron en el hecho de que la palabra «bueno» tiene diferentes significados en diferentes contextos. Si esto es correcto, un análisis general de las palabras «bueno» y «malo» poco puede contribuir a un esclarecimiento del significado específico que tienen cuando son empleadas en un contexto moral. Ciertamente, puede decirse que todos los enunciados simples en los que aparece la palabra «bueno» son enunciados valorativos, pero naturalmente con esta explicación no hacemos más que posponer el problema, ya que se puede dudar de que haya tantas clases de enunciados valorativos como significados de la palabra «bueno». Para algunas clases de enunciados valorativos puede existir una clase de fundamentación, para otros otra y para un tercer grupo ninguna.

Pero puede llevarse la crítica todavía más lejos. Por ejemplo, Habermas cree que las convicciones morales no se expresan en enunciados valorativos. Según Habermas, [el ámbito de] la moral no se expresa en oraciones evaluativas sino normativas. Lo característico de las oraciones morales no es la palabra «bueno», sino la palabra «deber». Por su parte, Hare comparte la vertiente positiva de esta concepción, pero no la negativa. Según Hare, las oraciones simples en las que aparece la palabra «bueno» y las oraciones simples en las que aparece la palabra «deber» son intertraducibles.[2] Creo que tanto Hare como Habermas están en un error. Sostengo, contra Habermas, que no es accidental que mu-

2. R. M. Hare, *The Language of Morals*, Oxford, 1952, cap. 12.

cho antes de la ética analítica toda la tradición de la filosofía moral haya estado orientada por ambas palabras «bueno» y «deber». Y contra Hare me parece claro que, aunque la mayoría de las oraciones valorativas pueden ser traducidas a oraciones normativas, lo contrario no es válido. Por ejemplo, las reglas del juego o las normas jurídicas no pueden ser traducidas a oraciones valorativas; esto parece mostrar que los enunciados normativos son de más diversas clases que los enunciados valorativos. Por consiguiente, se puede dirigir en este campo una crítica similar a la clarificación del concepto moral por el análisis de la palabra «deber» como a la orientación hacia el análisis general de la palabra «bueno».

Podría parecer atractivo llevar esta línea de crítica un paso más allá. Si ni la palabra «bueno» ni la palabra «deber» nos sirven para dilucidar lo que entendemos por moral, ¿no se sigue de ahí —podría preguntarse— que todo planteamiento semántico (la idea de que el significado de los juicios morales tiene que ser aclarado mediante un análisis semántico de ciertas expresiones) es incorrecto? Pero este paso no me parece justificado e, incluso, me parecería absurdo; tendría por consecuencia la renuncia, o cuando menos la no dilucidación del concepto de moral. No hay ninguna posibilidad de delimitar un concepto sino mediante el empleo de expresiones lingüísticas. Es cierto que no se puede definir las oraciones en que se expresan juicios morales mediante la presencia de palabras determinadas como «bueno» o «deber»; pero lo que cuenta en un análisis semántico no es la palabra, sino su empleo.

La consecuencia que podemos extraer de esto no es, por tanto, que hay que rechazar el planteamiento semántico sino, primero, que tiene que ser afinado; y, segundo, que en una clarificación conceptual es inevi-

table una cierta dosis de arbitrariedad. No hay una única respuesta a la pregunta por el criterio de los juicios morales. La pregunta no es: «¿qué es la Moral?», sino «¿qué entendemos por «moral»? Moral no es algo que se da objetivamente en todo lugar. Y como nuestra comprensión prefilosófica de «lo moral» es esencialmente vaga, no tiene perfiles definidos, es inevitable que siempre se dé algún aspecto de nuestra confusa comprensión prefilosófica de la «Moral» que equivoque de una u otra forma a nuestro concepto filosófico. El filosofar consiste en trazar líneas, en delimitaciones. La línea puede trazarse así o asá. Una vez la trazamos de una determinada forma, ponemos expresamente juntas algunas cosas y las separamos expresamente de otras con respecto a como estaban antes. Wittgenstein, en su *Cuaderno azul*, ha comparado este proceder con el orden de los libros de una biblioteca. Ningún orden es *el* orden correcto. Pero tenemos que poner algún orden en las cosas [objeto] de nuestra comprensión si queremos formular de ellas determinadas expresiones. Podemos comparar entonces un orden con otro y ver cuáles son las consecuencias, los beneficios y los perjuicios.

En nuestra interrogación semántica tenemos que considerar en particular tres palabras: primero la propia palabra «moral» y, en segundo lugar, determinados usos de las palabras «bueno» y «deber» a los que esta palabra parece remitir. El modo de utilización fáctica de la palabra «moral» no puede ser nuestro último criterio. Tanto desde el punto de vista sincrónico como sobre todo diacrónico, las palabras «moral» y «Moral» son polisémicas, y hubo épocas en las que no existieron de forma generalizada, a pesar de lo cual no querríamos decir que no se diese entonces lo que hoy queremos significar con un determinado sentido de estas

palabras. De lo que se trata en la fijación conceptual filosófica no es de calcar el modo de empleo fáctico de una palabra, sino hallar determinadas distinciones que, por cuanto podemos ver, son fundamentales desde el punto de vista antropológico (y esto quiere decir aquí más exactamente: sociológico, en la vida común intersubjetiva), por mucho que en el decurso de la historia se hayan articulado en diferentes palabras. Nuestra delimitación conceptual de la «Moral» debe ser suficientemente amplia para que posibilite un diálogo histórico acerca de las cuestiones morales a lo largo de las diversas épocas y para que una tal definición no prejuzgue las cuestiones sustantivas sobre qué normas morales son correctas. Por último, debe ser también lo suficientemente amplia para que pueda comprenderse por qué en diferentes épocas se articularon distinciones comparables.

Estamos por lo tanto ante una compleja tarea. Y resulta adicionalmente compleja por cuanto, inmersos en la tradición filosófica, no podemos partir de cero. Por eso tenemos que comprobar la cuestión de si en lo referente a convicciones morales se trata de un tipo de juicios de valor, como también la cuestión de si se trata de convicciones sobre un tipo de normas que no son traducibles a juicios de valor.

Empecemos por los juicios de valor, y por lo tanto por la palabra «bueno». Reflexionemos en primer lugar sobre cómo se emplea la palabra en contextos inequívocamente extramorales. La palabra se presenta rara vez como simple predicado, como cuando decimos lisa y llanamente de algo que es bueno o malo. Habitualmente la palabra se califica de determinados modos. Veamos algunas posibilidades.

Una primera posibilidad es la que podemos denominar modo de empleo hipotético: «*Si* quieres adelga-

zar», o mejor «*Para* adelgazar, es bueno comer menos». Se trata aquí de una conocida forma de hablar de buenos medios para un fin establecido de antemano.

Un segundo uso es aquel en el que decimos que algo es bueno para algo o para alguien: «Le ira bien casarse», «Sería bueno un poco de abono para las flores». En este segundo uso, que es detalladamente explicado por Von Wright, «bueno» significa lo mismo que «conveniente» y «malo», lo mismo que «perjudicial». Este uso de la palabra «bueno» está esencialmente ligado al uso adverbial de «bueno», en el cual decimos de algo que ello *le va bien*. Es conveniente no aquello que propicia un fin cualquiera de alguien, sino su mismo bienestar.

Un tercer uso de la palabra «bueno» es su uso como adjetivo atributivo, por ejemplo cuando hablamos de un buen cuchillo, un buen futbolista, un buen vino o una buena pieza musical. Los ejemplos muestran que no se trata de una clase unitaria. Pero para mis propósitos no es preciso explicar más detalladamente las subclases. Las más importantes son aquellas en las que se trata del grado de idoneidad de algo en la realización de su función característica (como en el caso del buen cuchillo y del buen futbolista) y también las que pueden designarse como de uso estético (como con respecto a un buen vino y una buena pieza de música).

¿Hasta dónde se llega entonces con estos tres casos en la dilucidación general de la palabra «bueno»? En todos estos casos, «bueno» parece remitir a una elección y una preferencia. El comparativo «mejor» parece tener prioridad con respecto al adjetivo positivo «bueno». Es un caso similar el de adjetivos como «largo» y «caliente». Siempre se trata de cosas que ordenamos en una escala. Algo es más largo que otra cosa cuando está más alto en la escala de longitud, y algo es mejor

que otra cosa cuando está más elevado en la escala de preferencia, y es bueno cuando está bastante alto en esta escala o más alto que la media.

Tenemos que distinguir ahora entre expresiones de preferencia subjetiva y objetiva. Cuando digo «Este cuchillo me gusta», expreso una preferencia subjetiva. Es característico de las expresiones de preferencia subjetiva que hay que designar un sujeto, que es quien prefiere. Cuando, por el contrario, decimos «Este cuchillo es bueno» o «mejor que aquél», no se presenta ningún sujeto de la preferencia. El carácter preferencial parece referirse al objeto en cuanto tal. ¿Cómo hemos de entender esto? ¿Por medio de qué se constituye una característica de preferencia objetiva? Una preferencia es objetiva cuando no es la sensación de una persona lo que decide sobre la preferencia, sino, por así decirlo, la misma cosa es la que reclama que sea preferida, con lo que se postula la preferencia para todas las personas en igual medida. En la mayoría de los casos esto adopta la forma de que la preferencia puede ser fundamentada. Con los motivos se justifican las expresiones valorativas y, con ello, la preferencia. Esto no parece ser así sólo en el caso de las expresiones de valor estéticas, que parecen rayar en una declaración de preferencia subjetiva. Afortunadamente, en el contexto actual no tengo que entrar detalladamente en esta dificultad.[3]

Lo esencial es que parece desprenderse una significación nuclear de la palabra «bueno», que esta palabra tiene no en todas pero sí en muchas de sus formas de uso. A resultas de esta significación nuclear, llamamos a algo «bueno» cuando creemos poder fundamentar

3. Véase Ernst Tugendhat, *Selbstbewusstsein und Selbstbestimmung*, p. 275.

que es preferible. Esta significación nuclear corresponde así siempre a un tipo de enunciado de deber, y esto quiere decir una norma. «Este cuchillo es mejor que aquel» implica pues «debería preferirse». Naturalmente, tiene mayor importancia dilucidar a continuación qué se entiende exactamente por «deber». En el caso actual parece significarse con ello lo siguiente: «sería irracional no preferirlo». Esta explicación me parece válida para todos los tipos de uso antes citado, a excepción del uso estético. En el caso en que «bueno» se utiliza atributivamente y se entiende de modo funcional, entendemos por «bueno» que cuando se quiere elegir algo de esta clase, por ejemplo un cuchillo, sería irracional no elegir éste. En el modo de uso hipotético, decimos explícitamente: si se quiere alcanzar X, es bueno —es decir: se debe, es decir: es racional— elegir Y. Algo similar sucede cuando se considera algo bueno para alguien. Se supone entonces que sería irracional que esa persona no lo eligiese. En todos estos casos, con el uso de la palabra «bueno» se pretende que hay razones para la preferencia. En el presente contexto puedo dejar abierta la cuestión de en qué consiste la fundamentación en cada uno de los diferentes casos, y hasta dónde llega, y sólo diré dogmáticamente que creo, primero, que la fundamentación no es nunca perfecta, que siempre queda un resto subjetivo, y, segundo, que, llegue hasta donde llegue la fundamentación, siempre puede ser reducida a expresiones empíricas.

Hay que recordar que todo enunciado valorativo de este tipo, en el que la valoración formula una pretensión de fundamentación, es traducible a un enunciado normativo, un correspondiente enunciado de deber. Para distinguir una norma así de otras normas, de las que voy a hablar en seguida, puede llamársela 'norma

racional', pues tiene el sentido descrito: «si no obras así, obras irracionalmente».

Ahora estamos ya en condiciones de poder plantear la pregunta de si hay también un tipo de uso específicamente moral de la palabra «bueno» y, si lo hay, si pertenece a la significación nuclear recién esbozada de esta palabra. Resulta natural identificar al tipo de uso moral como el tipo de uso de la palabra «bueno» en calidad de predicado simple, antes omitido.[4] La palabra «bueno» se predica de determinadas acciones. Cuando decimos de una acción que es buena (o, con más frecuencia, mala) sin más, y no con respecto a este o aquel fin o para esta o aquella persona, decimos, al parecer, que es moralmente buena (o mala). Podemos denominar a esto un uso absoluto de la palabra, con lo cual quiero significar, simplemente, sin una cualificación posterior. Ahora bien, ¿qué queremos decir cuando empleamos así esta palabra? ¿Queremos decir que sería

4. No es el único camino que puede describirse y que se ha descrito. Otra posibilidad sería considerar la significación moral original de «bueno» como atributiva, relacionada con la expresión de la clase «ser humano». Serían entonces buenas aquellas acciones realizadas por un ser humano bueno. Éste es el camino descrito por Aristóteles. Con ello ha dado este cuestionable segundo paso, el de entender la referencia a un «buen ser humano» análogamente a la de un «buen cuchillo», por lo tanto funcionalmente. Si se rechaza esto, se plantea en este enfoque la cuestión de a qué preferir racional se refiere esta consideración de un «buen ser humano»: si tengo que elegir entre diferentes seres humanos (¿para qué?), aquel que elegiría racionalmente es el más moral, y esto sería entonces una definición de lo moral [-mente bueno (T.)]. El enfoque aristotélico corresponde al uso moral de la palabra «bueno» en el griego prefilosófico. Pero los griegos tenían aún otra palabra prefilosófica, relacionada inmediatamente con las acciones morales, y que más o menos concuerda con nuestro «bueno» en el tipo de uso absoluto antes descrito: καλόν (bello).

de algún modo irracional no obrar así? A menudo se ha establecido esta vinculación entre razón y moral, pero me parece claro que esta concepción es contraria a nuestra comprensión normal de lo moral [-mente bueno (T.)]. Hacer algo inmoral no significa comportarse irracionalmente, y creo que no deberíamos aceptar ninguna explicación de la moralidad que tuviese esta implausible consecuencia.

Es verdad que ciertos filósofos, todos ellos anteriores a Kant, han supuesto, al concebir una semejante vinculación entre la acción moral y la racional, un *particular* concepto de razón. Pero en sentido estricto se trata de una equivocación por uso incorrecto de una palabra. La referencia habitual a una acción y preferencia razonada y racional indica que se trata de una acción o preferencia que es de tal modo fundamentable que de quien no actúa de tal modo decimos que se conduce irrazonable, irracionalmente. A quien obra inmoralmente se le podría objetar no que obra irracionalmente, sino contra las mejores razones. «Lo que haces —podría objetarse a alguien que obra inmoralmente— va en tu propio interés bien entendido y, por lo tanto, es racional, pero ¿cómo puedes razonar el actuar así?» Pienso sin embargo que aquí se ha interpretado la forma de hablar cotidiana de modo impreciso. Se preguntaría al interpelado no «¿cómo puedes fundamentar el obrar de este modo?», sino «¿cómo puedes fundamentar tener que obrar de este modo?». Pero esto significa que en esta cuestión de fundamentación está ya implícita —con respecto al «tener»— una relación a algo normativo; lo normativo constituye por consiguiente el punto de referencia de la fundamentación y no se constituye, como en las normas racionales, por medio de éstas.

Supuesto, por lo tanto, que podemos identificar

realmente el modo de uso «absoluto» con el moral, este modo de uso de la palabra no entra en el ámbito de significación nuclear antes esbozada, pues las normas morales no son «normas racionales». Pero sin este presupuesto llegamos al mismo resultado. Pues, entendamos como entendamos el modo de uso absoluto de la palabra «bueno», si por «bueno» entendemos también aquí —donde esta palabra no está cualificada es decir, no es relativa a algo— que está fundamentado preferir lo que es bueno, ello significaría que existe una fundamentabilidad práctica que no es relativa a algo. Podría quizá decirse: ¿por qué no? ¿Por qué no debe ser posible fundamentar una acción o elección no sólo por relación a un fin, y no sólo por relación al bien de un ser, sino fundamentarla sin más? Sobre ello sólo puedo decir que no parece nada claro qué puede significar que una acción se fundamente sin más, y no por relación a algo, o que es racional querer algo, no con respecto a un fin y no con respecto al bien de un ser, sino racional sin más. Según puedo ver se trata de una idea absurda.[5]

Así pues, dejo también en el aire este modo de uso de la palabra «bueno». Cuando una proposición valorativa no corresponde a una norma racional, no resulta nada claro qué puede significarse con ella. En vez de seguir ensayando, propongo considerar fracasado este primer intento de delimitar el concepto de lo moral, desde la perspectiva de las proposiciones de valor.

Remitámonos, por lo tanto, a la otra posibilidad, a saber, que las convicciones morales están relacionadas primariamente con normas, lo cual sólo puede signifi-

5. Yo mismo la he presentado en mis *Vorlesungen zur Einführung in die sprachanalytische Philosophie* (*Lecciones introductorias a la filosofía analítica del lenguaje*), 7.ª lección.

car naturalmente: con normas que no son traducibles en juicios de valor, y por consiguiente con normas que no son normas racionales (en el sentido antes definido). Tendría que mostrarse si quizá mediante este rodeo puede alcanzarse aún una comprensión del modo de uso absoluto de «bueno», que tan difícil resultó obtener de forma directa.

En términos muy generales, puede entenderse por normas indicaciones generales de acción [*allgemeine Handlungsanweisungen*] formuladas en proposiciones (a diferencia de aquellas reglas de acción que sólo son reconocibles por el hecho de que determinadas acciones —por ejemplo, pasos de baile o secuencias de sonidos— son valoradas como correctas o incorrectas). El obrar regido por normas es por consiguiente un tipo de conducta reglamentada, que debe distinguirse de la conducta meramente conforme a reglas. Una indicación de acción formulada en una proposición puede comprenderse como un imperativo general. En filosofía se acostumbra a calificar con frecuencia a las normas como proposiciones de deber, pero «deber» es la palabra normativa débil; la palabra que utilizamos normalmente en el discurso normativo de la vida real es la palabra normativa fuerte «tener que» (o, por negación, «no tener que» o «no poder»).[6] Es constitutivo de todo obrar normativo, y en general de todo obrar reglamentado, que es valorable como correcto o incorrecto con relación a algo, y en cada caso, según aquello en relación a lo cual es criticado se distingue así el sentido del «debe» (o «tiene que»). El mayor peli-

6. Tanto «deber» como «tener que» tienen también un uso teórico para designar regularidades fácticas. Para una detallada presentación de estos contextos, véase John Leslie Mackie, *Ethics, Inventing Right and Wrong*, Harmondsworth, 1977, cap. 3.2.

gro de desconcierto que amenaza a la filosofía moral consiste en dejar en estado nebuloso el sentido del «debe».

Para comprender qué se entiende en cada caso por el «debe», creo que tenemos que preguntarnos siempre: ¿qué es lo que pasa (o cuál es la consecuencia) cuando no se hace lo que se debe? Ya lo hemos visto con respecto a las normas racionales; en éstas significa «debe preferirse», y si no se prefiere, se actúa irracionalmente. Como hemos visto que el sentido del «debe» moral no es el de estas normas racionales, resulta natural buscarlo en el ámbito de aquella clase de normas en las que lo que pasa, cuando no se actúa así, consiste en una sanción por parte de los demás (o de los dioses). Se entiende aquí por sanción toda la gama de reacciones negativas que van desde la desafección, la censura y el menosprecio hasta la imposición de sanciones externas. Las normas así definidas las denominaré normas sociales. Las normas sociales son constitutivas en cada caso de una praxis intersubjetiva.

Una semejante praxis intersubjetiva puede ser o bien una praxis delimitada, en la que un individuo entra libremente y a la que libremente puede renunciar, la forma más pura de cuya praxis, no contaminada con la otra, es la de los juegos sociales, como por ejemplo el fútbol, el ajedrez o el juego de canicas, tan detalladamente estudiado por Piaget. O bien la praxis intersubjetiva se refiere al conjunto de la vida social de una comunidad, y las normas que rigen esta praxis son aquellas que pueden denominarse, en sentido estricto, normas sociales:[7] las costumbres y las normas jurídicas.

7. Algunos autores utilizan incluso el término general de «norma» sólo en este sentido; estas cuestiones terminológicas son, obviamente, de poca importancia.

En tanto en cuanto en cada caso lo que se reglamenta es la totalidad de una praxis, bien restringida o bien global, las normas sociales se amoldan a sistemas, por ejemplo un sistema jurídico o la totalidad de reglas que definen un juego. Así, un sistema normativo consiste en parte en prescripciones, pero también esencialmente en prohibiciones, y lo que no está prohibido ni ordenado, está permitido: definen por lo tanto las condiciones restrictivas en las que el participante de ese juego o sociedad puede hacer lo que quiere, lo que también significa, obviamente, aquello que dado el caso quiere racionalmente (y, en este sentido, debe). De este modo hay, por ejemplo en el ajedrez, dos enunciados de deber totalmente diferentes. Por una parte, están los enunciados normativos, que atañen a las reglas del juego, como cuando digo a un principiante «así no puedes (no está permitido) mover el alfil». Por otra parte, nos encontramos con otros enunciados normativos como «así no puedes mover», con lo cual queremos decir: «este sería para ti un mal avance». En el segundo enunciado está implícita la idea de «de otro modo obras irracionalmente», mientras que en el primero la de «de otro modo transgredes una regla vigente».

Es característico de las normas sociales, y sólo de ellas, que «valen». ¿Qué quiere decir esto? Cuando alguien transgrede las reglas válidas de un juego, ello tiene por consecuencia que los demás no quieran seguir jugando con él; cuando transgrede las costumbres vigentes, ello tiene la consecuencia de que es más o menos marginado o menospreciado; cuando lo que transgrede es el derecho vigente, o comete pequeñas infracciones de juego, tiene que acarrear determinadas sanciones preestablecidas. No sería comprensible qué significa el «deber»/«tener que» y el «valer» de la nor-

ma social sin la sanción correspondiente a la ejecución de lo no debido.

Estamos buscando una determinación conceptual de los juicios morales. Una vez se ha mostrado que la moral no está en el ámbito de las normas racionales, resulta natural buscarla en el ámbito de las normas sociales. Pero la moral se expresa característicamente en juicios, y los juicios reclaman una pretensión de fundamentación. Si el intento de encontrar el lugar de las normas morales en el ámbito de las normas sociales ha de tener perspectivas de éxito, debería poderse mostrar que las normas sociales pueden ser referidas de algún modo a la fundamentabilidad.

Con relación a esto se aprecia en seguida un neto contraste entre las reglas de juego, por un lado, y las normas sociales en sentido estricto, por otro. En el ajedrez, por ejemplo, es absurdo pedir una justificación o fundamentación de las reglas del juego. Las reglas del juego no parecen ser susceptibles de fundamentación, y por cierto no porque no la necesiten, sino porque queda al arbitrio del individuo el querer o no querer tomar parte en esta praxis intersubjetiva. (Dicho sea de paso, una consecuencia de la voluntariedad de los juegos es que sus reglas son siempre justas; no benefician a un participante frente a otro, por el simple motivo de que nadie querría participar en un juego con semejantes reglas.) Por el contrario, los individuos están sometidos a los usos y normas jurídicas de su sistema social, tanto si lo quieren como si no. Y como no se trata de leyes naturales sino de sanciones humanas, se plantea la cuestión de cómo hay que justificar una tal restricción de la libertad del individuo contra su voluntad. Por consiguiente, las costumbres y normas jurídicas están *necesitadas de fundamentación* [*begründungsbedurftig*].

Resulta natural entonces designar como principios morales a aquellos con cuya ayuda pueden fundamentarse los usos y normas jurídicas, de lo cual se seguiría asimismo que, por su parte, también los usos y normas jurídicas, cuando son así fundamentados, pueden denominarse morales. Expresado en términos más simples: resultaría que aquellas normas sociales que se considera fundamentadas se denominarían morales, con lo cual estaría definida la palabra «moral».

Esto es sin embargo precipitado. En primer lugar deberíamos interrogarnos qué puede significar en general fundamentar una norma social. Para apreciar correctamente la dificultad de esta cuestión, es conveniente formularla aun en términos más generales: ¿qué puede significar, en general, fundamentar una norma de cualquier tipo? Una norma es un imperativo generalizado, y un imperativo es un tipo de oración que, por contraposición a una oración predicativa, no se apoya en una fundamentación; no puede verse cómo puede concebirse aquí la pregunta por una fundamentación. Esta dificultad podría obviarse sin embargo fácilmente porque, como hemos visto, hay un tipo de oraciones de deber, a saber las normas racionales, que es realmente fundamentable, lo cual se explica con facilidad por el hecho de que una oración de este tipo es traducible a una oración predicativa. En este último caso se fundamenta que a algo le corresponde el predicado «bueno» en relación a algo. En este predicado resulta también inteligible *qué* es lo que en este caso se fundamenta. Siempre que una oración es fundamentable, tiene que formular de por sí una pretensión de fundamentación. Pero, ¿sobre qué debe formular una norma jurídica o un uso una pretensión de fundamentación?

Puede resultar instructivo entrar brevemente en

una concepción en la que se aborda este problema, la concepción de Habermas. Habermas parte de la consideración correcta de que es un hecho histórico-sociológico que en las costumbres y normas jurídicas se interroga y siempre se ha interrogado acerca de su fundamentación. Pero constatar este hecho no significa comprenderlo. De una norma social no puede decirse de igual modo que de una oración predicativa que es directamente esta misma oración la que puede ser fundamentada. Habermas intenta concebirla igualmente por cuanto cree poder distinguir entre diferentes «pretensiones de fundamentación», la de la verdad en las oraciones predicativas y la 'correctitud' en las normas. Pero tanto la referencia a la «validez» como a la «correctitud» no están en este contexto nada claras. Cuando se habla de «validez» en relación a normas, se alude normalmente a su validez positiva (por ejemplo, en un orden jurídico). Y cuando se habla de «correctitud» en relación a normas, se alude a la rectitud de una acción en relación a una norma. No resulta nada claro qué puede significar el decir de una norma que vale absolutamente o que ella misma es correcta. La respuesta de Habermas a la pregunta de qué es lo que fundamentamos al fundamentar una norma, sería: que vale o que es correcta, pero esta explicación es vacía, en tanto no puedan ser explicadas las palabras «validez» y «correctitud». La supuesta analogía con la verdad de las oraciones predicativas no se sostiene. Ciertamente podría decirse que lo que fundamentamos, al fundamentar una proposición predicativa, es que es verdadera, pero esto no significa nada más que fundamentar la propia predicación (aquí no necesitamos la palabra «verdadero»). A la pregunta de qué es lo que entiende por la correctitud de una norma, Habermas sólo puede contestar: que está fundamentada. Esto es

sin embargo circular, pues la pregunta era precisamente dónde está la fundamentabilidad.

La idea de que una norma como tal pudiese ser fundamentable, igual que una predicación como tal es fundamentable, resulta por lo tanto errónea. El único sentido posible que puede tener una fundamentación de una norma social es que está fundamentado que la norma tiene una cualidad determinada. Así debería atribuirse según Habermas a una norma la cualidad de la correctitud, lo que es por supuesto una pseudocualidad. La idea de que lo que hay que fundamentar aquí es una oración predicativa, en la que se atribuye a una norma una cualidad, pierde la posible apariencia de artificialidad tan pronto como se piensa en el caso particular de la pregunta por la legitimidad de las normas jurídicas, en la que resulta totalmente natural decir: lo que hay que fundamentar aquí es que las normas son justas. Aquí es por consiguiente la justicia la cualidad buscada.

Pero éste es un caso especial. Para aproximarse al caso general, resulta natural partir precisamente del factor que hace que las normas sociales en sentido estricto, por contraposición a las reglas del juego, estén necesitadas de fundamentación. Pues aquello que hay que fundamentar en una norma tendría que atañer precisamente a aquello que la hace necesitada de fundamentación. Ya hemos visto que las reglas de juego no son susceptibles de fundamentación porque no están necesitadas de ella, y que no lo están porque en la praxis intersubjetiva de un juego se entra libremente y con carácter temporalmente limitado. Por el contrario, las normas sociales en sentido estricto mostraron estar necesitadas de fundamentación, porque los individuos están sometidos a ellas de una vez por todas mediante sanciones. Exigen al individuo posponer sus

propios intereses o limitar su libertad en virtud de ellas. Hay entonces dos posibilidades, que el individuo sienta esta compulsión como mera compulsión o la conciba como fundamentada, y en esta medida incorpore la compulsión en el ámbito de su propia libertad.[8] En este último caso, el individuo tendría con respecto a las normas sociales precisamente aquello que tiene a su favor el sistema normativo de un juego por contraposición al sistema normativo de la realidad, a saber, la voluntariedad de la sumisión. La fundamentación tendría que consistir, por tanto, en que el individuo se convenciese de que, si tuviera la posibilidad de decidir con respecto a determinadas partes o al conjunto del sistema normativo en el que se encuentra inmerso no libremente y para toda su vida, lo habría elegido con igual libertad con la que puede elegir en todo momento entrar en un juego.

Pueden pensarse diversos motivos (volveré a ello más adelante) que hacen que un individuo asienta a un sistema de normas sociales, lo cual significa precisamente esto: le permiten afirmar que también se hubiera sometido o estaría sometido libremente. Pero este asentimiento es racional si tiene por motivo que el individuo llega a concebir que la validez de la nor-

8. Hay que hacer dos observaciones con respecto a esta afirmación: en primer lugar, muestra lo falsa que es la idea de que la explicación del «debe» de las normas sociales representa una alternativa con respecto a las sanciones y con respecto a la fundamentación, pues esta última se basa, por su mismo sentido, en la primera: lo que hay que fundamentar es precisamente el carácter sancionador de las normas. En segundo lugar, podría dudarse de si aún tiene sentido hablar de compulsión cuando se asume algo libremente. Para aclararlo puede ser de utilidad aquí la comparación con el juego, como puede verse en la siguiente afirmación del texto.

ma (o de todo un sistema normativo) es *buena para él*, o que es mejor para él que todos estén sometidos, que no que nadie esté sometido a ella. Éste es el caso cuando el beneficio que espera la persona de que todos los demás tengan que seguir esta regla es mayor que el perjuicio derivado de tener que cumplirla. Puede pensarse en una escala en la que un individuo puede ordenar los sistemas normativos según el criterio «mejor para mí».

Con ello hemos hallado ya un primer predicado con el que pueden ser valorados las normas y sistemas normativos. Por supuesto, este predicado «bueno para uno» no posibilita aún una fundamentación de la validez de la norma como tal, sino sólo una fundamentación de que un individuo particular o una clase social de individuos asientan a la norma, y por cierto racionalmente. Pero como entonces existe la misma necesidad de fundamentación para un individuo (o para una clase de individuos) que para todos los individuos, resulta como siguiente paso obvio la cuestión siguiente: ¿Es la norma buena para todos (y para todos de igual modo)? Resulta entonces evidente decir: precisamente es esto aquello a lo que se alude con el modo de uso absoluto de la palabra «bueno». Esto tendría entonces también por consecuencia que el modo de uso de la palabra «justo» se puede integrar de forma obvia en el de la palabra «bueno». Pues es entonces natural decir que el aspecto específico de que la norma no es sólo buena para todos, sino buena para todos de igual modo, de que por consiguiente es no partidista (y se trata de un aspecto sólo relevante para determinadas normas sociales), se expresa especialmente con la palabra «justo».

El resultado sería, por lo tanto, el siguiente: aquello que fundamentamos al fundamentar las normas es

que son buenas, lo que quiere decir que son buenas para todos en igual medida.

Se puede entender entonces igualmente que, cuando fundamentamos una norma social, esto tiene el sentido de que fundamentamos que le corresponde una determinada cualidad o, en términos lingüísticos, que le corresponde un determinado predicado. En cualquier caso ha resultado ahora claro en un ejemplo que esto tiene sentido, y qué sentido tiene, mientras que la alternativa —que puede fundamentarse una norma en y por sí— ha mostrado ser vacía. Podría querer preguntarse entonces: primero, ¿es forzoso que el predicado de fundamentación sea precisamente la palabra «bueno»?; y, segundo, si es ya la palabra «bueno», ¿tiene que entenderse esta palabra en el sentido de «bueno para todos»?

Creo incluso que pueden unirse ambas preguntas. En primer lugar, hubo y hay culturas que fundamentaron de tal modo sus normas que reclamaron un origen divino para ellas. El predicado relevante no era entonces «bueno», sino acaso «santo» o «querido por Dios». En segundo lugar, aun cuando el predicado de fundamentación sea la palabra «bueno», no es en modo alguno necesario entenderlo precisamente tal y como he propuesto. Por contraposición al modo de uso relativo, el modo de uso absoluto de la palabra «bueno» no parece tener en el lenguaje una significación unívoca.

Concibo en particular otras dos posibilidades en que puede entenderse en nuestro contexto, y de hecho se ha entendido, esta palabra. La primera es aquella concepción en la que una norma está fundamentada, no cuando es buena para todo individuo, sino cuando es buena para la comunidad como un todo. También aquí se definiría por lo tanto el modo de uso absoluto mediante el modo de uso relativo en el sentido de con-

veniencia, sólo que conveniencia ahora no con respecto a todos los individuos particulares, sino entendida con respecto al grupo, concebido como un único ser. Esto presupone una concepción holista de la relación de los individuos con el grupo; los individuos existen para la comunidad, y no al revés. La comunidad se concibe según el modelo de un organismo, y los individuos como sus órganos. Esta concepción de la moral no es obviamente extensible a una moral universal, como es aquella en la que se concibe a los individuos como último e igual punto de referencia. A esta comprensión organicista, ligada al interés común, del uso absoluto de la palabra «bueno» va asociada también una correspondiente comprensión de la palabra «justo». La antigua concepción de la justicia expresada en el principio de «a cada cual lo suyo» puede entenderse en el sentido de la imparcialidad, pero también puede ser concebida de modo organicista: cada uno no cuenta entonces igual, sino que merece lo que le corresponde de acuerdo con su posición en el sistema jerárquico de la comunidad.

Una tercera concepción del modo de uso absoluto de la palabra «bueno» es la utilitarista. De acuerdo con la concepción del utilitarismo de la regla, un sistema normativo A es mejor que un sistema normativo B, cuando A produce más bienestar que B, con lo que resulta indiferente cómo se distribuye el bienestar. En oposición a la concepción organicista, en la concepción utilitarista los individuos constituyen el punto de referencia último, y el bienestar de un individuo no cuenta más que el de otro, pero lo que cuenta es el conjunto total de bienestar, y no su igual distribución.

Lo que distingue a «bueno» en el sentido de «en igual medida para todos» tanto de la concepción organicista como de la utilitarista, es que se dispone en dos

niveles: el primero consiste en la interrogación de todo individuo, si puede asentir racionalmente al sistema normativo, y el segundo en la universalización de este asentimiento.

En la tercera lección intentaré mostrar en qué medida esta concepción, según la cual las normas sociales están fundamentadas cuando son buenas en este sentido particular de «ser buenas en igual medida para todos», es distinta de las demás concepciones. Lo que me parece esencial en el contexto actual, referente ante todo a la delimitación de un *concepto* de la moral, es que este concepto resultaría impropiamente restringido si quisiéramos poder hablar sólo de una fundamentación de las normas sociales como morales, cuando están bien fundamentadas. Si se quisiera restringir de este modo el concepto de fundamentación moral, limitando la multiplicidad de posibles predicados de fundamentación («otorgado por Dios», «bueno$_1$», «bueno$_2$», etc.) a uno sólo, se desecharía con ello la confrontación entre estas diferentes concepciones de fundamentación moral, o bien se prejuzgaría con un golpe de fuerza semántico. Mediante la equiparación de «moral» con un determinado predicado de fundamentación favorito, por ejemplo «bueno$_1$», habríamos excluido simplemente todas las demás concepciones de fundamentación a partir de una definición. De este modo, se falsearía una cuestión sustancial en una cuestión de palabras. Aquí chocamos por consiguiente con el justificado núcleo del recelo contra los métodos semánticos. Pero naturalmente no atañe al método semántico como tal, sino a una aplicación defectuosa. Existe un criterio claro a condición del cual estaría justificada semejante restricción del concepto de fundamentación moral. Sería cuando no se pudiera formular una confrontación significativa entre las diversas concepciones

antes citadas, que se articulan en cada caso en un correspondiente predicado de fundamentación. O bien tienen que ver con diferentes cosas, con lo que no pueden chocar entre sí; o bien tienen que ver con diferentes concepciones de una misma cosa. Pero entonces es obligatorio captar esta cosa bajo un concepto unitario.

Resulta natural utilizar la palabra «moral» como expresión de este concepto unitario. Pero incluso en esta palabra en cuanto tal no hay nada natural. No en todas las culturas existió una palabra semejante. Lo que sí existió en todas las culturas, y sobre lo que hubo diferentes concepciones, es la pretensión de que las normas sociales (o una parte de ellas) están fundamentadas. Por ello resulta natural retener con la palabra «moral» precisamente esta cualidad de las normas —el hecho de que se las considera fundamentadas, se entienda como se entienda esto más detalladamente, es decir, en qué predicado de fundamentación se articule esta pretensión de fundamentación.

Pero, si hay diversos predicados de fundamentación posibles, ¿en qué consiste entonces el sentido unitario de la referencia a una fundamentación de las normas sociales? Ya he mostrado que no es significativo hablar de una fundamentación de una norma social en y para sí, tal y como se puede fundamentar una oración enunciativa en y para sí. Si aquello que fundamentamos es un producto lingüístico, puede tratarse siempre sólo de una oración enunciativa; por ello, aquello que podemos fundamentar, cuando fundamentamos una norma, es sólo que le corresponde un determinado predicado. Pero después de constatar que se proponen varios predicados de este tipo, se podría intentar buscar un superpredicado en el que se expresase la unicidad de aquello que tenemos que fundamentar. Mas esto sería

tan erróneo como la idea de que se puede fundamentar en y por sí una norma. La pregunta de qué es, pues, en general lo que fundamentamos cuando fundamentamos una norma social, sólo se puede responder oportunamente reflexionando de nuevo sobre qué era lo que hacía que las normas sociales estuvieran necesitadas de fundamentación. Ello era la circunstancia de que estamos sometidos a estas normas y a sus sanciones, queramos o no, y que por nuestra parte exigimos a los demás (por ejemplo, a nuestros hijos) la misma restricción. Esta restricción de nuestro obrar y la simultánea exigencia de que los demás limiten igualmente su conducta, es lo que hay que fundamentar.

Aquello que fundamentamos por consiguiente cuando fundamentamos una norma (o, mejor dicho, la validez y, por tanto, la sanción de la norma) no es por consiguiente ni esta norma como tal ni una oración enunciativa en la que predicamos algo de una norma, sino que pensamos tener un motivo para sometemos libremente a la norma y su sanción. El sentido de hablar de una fundamentación no es por consiguiente en última instancia una fundamentación *de* (un enunciado), sino una fundamentación *para* (un obrar): la fundamentación es una fundamentación para participar en la praxis intersubjetiva que se define mediante un sistema normativo. Un individuo tiene para las diferentes acciones y complejos de acción de su vida diferentes razones (motivos), y la sociedad o bien sus congéneres le exigen conducir estas acciones, sean cuales sean sus motivos, bajo determinadas condiciones restrictivas. Puede preguntar: ¿por qué? Se le responderá que porque a estas normas les corresponde una cualidad distintiva; obviamente, esta respuesta debe tener la fuerza de constituir para el individuo un motivo para que éste acepte libremente esta restricción de su

obrar o, más allá de éste, de toda la praxis intersubjetiva determinada por las normas.

Naturalmente, un tal asentimiento sobre la base de la citada cualidad distintiva guardará relación con la autocomprensión del individuo en cada caso: si se concibe como hijo de Dios, la circunstancia de que las normas son queridas por Dios es el motivo para someterse a ellas y para exigir esta sumisión a los demás; si se concibe como miembro de una comunidad entendida a modo organicista, el hecho de que las normas sean buenas para la comunidad es motivo para someterse a ellas y exigir lo mismo a los demás miembros de la comunidad, etc. Con ello resulta ahora claro dónde radica el sentido unitario de la fundamentación de las normas y, al mismo tiempo también, de qué forma la fundamentación del *enunciado* de que a una norma le corresponde un determinado predicado de fundamentación, está relacionada con el intrínseco sentido *práctico* de la fundamentación de las normas sociales —como una fundamentación del ingreso en una praxis intersubjetiva.

Ahora se ve con claridad por qué la pregunta por el verdadero significado de la palabra «bueno», que ha desempeñado un papel tan considerable en la ética analítica clásica, es tan poco significativa para la cuestión de la fundamentación de las normas. No sólo no existe un firme significado del uso absoluto de esta palabra, sino que tampoco depende de esta cuestión. La cuestión no es cuál es la acepción más correcta de esta palabra, sino qué respuesta más idónea contiene a la pregunta por el motivo que podemos tener para someternos a las normas.

Con ello hemos alcanzado el límite de la dilucidación conceptual previa, es decir del acceso semántico al problema de la moral. La sospecha de que las refle-

xiones semánticas no son idóneas para la fundamentación de nuestros juicios morales ha resultado correcta, pero estas reflexiones eran necesarias para llegar a una comprensión correcta de qué conviene en general a la fundamentación.

Nuestra pregunta de partida era cómo se puede delimitar el concepto de juicios morales y en qué radica su fundamentabilidad. La respuesta es que los juicios morales remiten a normas sociales, y por cierto a normas que formulan una pretensión de fundamentación, pretensión que ha mostrado tener dos niveles: en un primer plano, por así decirlo interno, consiste en la pretensión de que a una norma le corresponde aquel predicado que expresa la cualidad considerada esencial para la moralidad, y es este predicado el que se emplea en las expresiones en las cuales se expresan juicios morales. En un segundo plano, por así decirlo externo, la fundamentación consistía en la prueba de que la pertinencia de este predicado constituye un motivo —y, eventualmente, un predicado mejor que los demás— para someterse a un sistema normativo así caracterizado y exigir esta sumisión también a los demás.

El primer plano parece comparativamente trivial, pues en él no se trata todavía de la cuestión de si las normas son fundamentables en última instancia. No obstante, en la segunda lección mostraré que también esta fundamentación inmanente tiene en cada caso su propia problemática, si el predicado de fundamentación es de tal tipo que no existe ningún procedimiento trivial para aplicarlo. Esto vale en particular en el caso de que el predicado de fundamentación es el que yo he elegido, a saber, «bueno en igual medida para todos». Sólo en la tercera lección pasaré al segundo plano de fundamentación y abordaré la cuestión de cómo se puede pensar una confrontación entre los diferentes

predicados de fundamentación, es decir, de si existe un patrón de algún modo absoluto y, por tanto, una fundamentación última.

2. ¿SE PUEDE APRENDER MORALMENTE A PARTIR DE LA EXPERIENCIA?

La primera lección tuvo por resultado que la pregunta por la fundamentación de nuestras convicciones morales se plantea en dos planos. El primero atañe a la fundamentación de normas morales bajo la presuposición de un determinado predicado de fundamentación moral, y puede decirse también: bajo la presuposición de una determinada concepción de la moral. El segundo plano se refiere a la fundamentación de esta misma concepción, o lo que es lo mismo a la confrontación entre las diferentes concepciones, y con ello al mismo tiempo a la pregunta por la fundamentación de un punto de vista moral en general.

Ambas problemáticas suelen abordarse con un planteamiento basado en la alternativa verdadero o falso. Se intenta probar que una concepción moral es la verdadera y las otras falsas; igualmente, en la fundamentación de normas en el seno de una concepción se intenta probar que hay un procedimiento de fundamentación definitivo, y si no se consigue —se dice— no se puede hablar de fundamentación en general. Frente a esta concepción, quiero intentar la consideración de ambas problemáticas desde un punto de vista comparativo. Con ello quiero indicar que quizás hay sólo mejores y peores fundamentaciones, y que si no hay una fundamentación última, resulta igualmente significativa una relación con normas relativamente funda-

da, esencialmente diferente de un puro decisionismo.

Ya se ha mostrado en la primera lección que no tiene sentido hablar de un verdadero principio moral, dado en sí de algún modo y del que derive un «debe» [*muss*] absoluto y con un sentido igualmente inexplicable. De existir semejante principio, tendría que tener un carácter sintético *a priori*, pues, en primer lugar, con seguridad semejante principio no podría ser empírico y, en segundo lugar, resulta claro que un principio *a priori* puramente analítico no podría contener un mandato sustancial. Por este motivo también Kant ha creído que su imperativo categórico —un principio que está aproximadamente en la línea de la explicación de la palabra «bueno» [*gut*] de la que soy partidario— es un principio sintético *a priori*. Kant ha considerado justamente que a partir de una mera explicación analítica de esta (o de cualquier otra) palabra (y puede decirse que toda la segunda sección de la *Fundamentación de la metafísica de las costumbres* no es otra cosa que esta explicación) no puede seguirse que debamos querer realmente las normas supeditadas a este principio. Por este motivo recurrió a la idea de un *a priori* sintético. No puedo entrar aquí en la problemática de este concepto. En la actualidad, casi todos están de acuerdo en que se trata de la cuadratura del círculo. Hoy sólo podemos comprender y explicar en términos históricos esta empresa kantiana de sustituir, tanto en la filosofía teórica como en la práctica, un «debe» [*muss*] fundado en una base trascendente y que se había perdido por un «debe» fundado en la esencia del sujeto. Pero si no puede fundarse absolutamente un principio moral ni objetivo ni subjetivo, subsiste la posibilidad de una fundamentación relativa o, por utilizar una formulación más cauta y negativa, la posibilidad de mostrar que está menos expuesto a contraargumentos que

otros principios, y por tanto la posibilidad de un no relativismo en sentido comparativo.

Esta cuestión voy a abordarla en la próxima lección. Ahora voy a entrar con el mismo espíritu en el problema de la fundamentación interna en el seno de una concepción moral ya dada. De hecho, esta aproximación comparativa es la habitual en la fundamentación de una teoría empírica. Tampoco aquí se trata de si la teoría es verdadera o falsa, sino de si una teoría es superior a la otra, y una vez hemos creído haberlo mostrado, todavía no damos a esta teoría un estatus absoluto, sino que esperamos que, por su parte, sea sustituida en el futuro por otra teoría mejor fundada. ¿Cómo puede parecer erróneo concebir de forma similar la fundamentación de normas morales? La respuesta parece clara: una teoría científica es empírica; son nuevas experiencias las que la debilitan y suscitan la necesidad de una teoría mejor, más completa. Por el contrario, las convicciones morales, sea cual sea el predicado de fundamentación con que se articulen, no atañen a hechos empíricos y por eso no parecen depender tampoco de la experiencia. Formulan una pretensión no empírica y, en este sentido, absoluta.

Por otra parte, parece igualmente natural decir que podemos aprender moralmente a partir de la experiencia, tanto individual como colectivamente. Se dice así, por ejemplo, que una persona puede adquirir, sobre la base de muchos años de experiencia, una mayor sabiduría moral. Y se imagina también que, por ejemplo, una nación pueda conocer, en el curso de su historia, un aprendizaje moral, por ejemplo con respecto a la sabiduría de su constitución. Pero para que estas ideas puedan tener sentido, la estructura de estos procesos de aprendizaje debe ser evidentemente de otro tipo que la estructura de lo que se conoce normalmente como

un proceso de aprendizaje, un proceso que se remite a hechos y que consiste en aprender o controlar regularidades empíricas y, eventualmente, causales.

Ésta es la cuestión que quiero analizar en esta lección: ¿existen procesos de aprendizaje moral, por lo tanto procesos en los cuales tanto los individuos como los grupos sociales pueden no sólo cambiar sus convicciones morales, sino también cambiarlas de tal modo que arriben a concepciones que (al menos según su propia concepción) sean mejores o más correctas que las que tenían anteriormente?; y, si los hay, ¿cómo pueden comprenderse estructuralmente semejantes procesos de aprendizaje práctico-moral por contraposición a los procesos de aprendizaje teóricos?

Naturalmente, lo que llamé pretensión de absolutidad de los juicios morales parece contradecir la posibilidad de una experiencia moral. Por este motivo resulta natural considerar ilusoria la opinión de los sujetos históricos convencidos de realizar experiencias morales. En verdad —se dice— ni en el caso del individuo ni en el caso de la sociedad se trata de un proceso de aprendizaje, sino meramente de un cambio de una convicción moral a otra. Si se entiende por experiencia histórica el ocuparse teóricamente de los cambios históricos, parece que la experiencia histórica de la historia moral da por resultado que no existe una experiencia específicamente moral.

Pero obviamente, este tipo de consideración historicista tiene consecuencias ulteriores. Lo que parece ilusorio en esta perspectiva no es sólo la posibilidad de un proceso de aprendizaje moral, sino también toda convicción moral en y por sí, la pretensión absoluta de estas convicciones. La experiencia teórica de la historia de la moral parece conducir al relativismo, y por tanto al desenmascaramiento de la moral en general.

En primer lugar, el relativismo resulta ya de la pura observación de la multiplicidad de convicciones morales recíprocamente contradictorias que hallamos en la historia y cada una de las cuales presenta una pretensión absoluta. Pero el relativismo histórico adquiere aún otra cualidad, en la cual alcanza su auténtico efecto de desenmascaramiento, si consigue reconducir a otra cosa las diferentes convicciones morales mediante explicaciones causales. Sólo entonces tenemos no sólo una pluralidad de términos que se ponen recíprocamente en cuestión, sino una auténtica relatividad en tanto en cuanto se muestra que aquello que parece tener un sentido absoluto desde el punto de vista interno de los agentes, es relativo con respecto a determinadas condiciones, por ejemplo, las condiciones socioeconómicas. Y lo que desde el punto de vista interno parecen procesos de aprendizaje moral, se entiende ahora como un mero cambio, condicionado por un cambio del ambiente. Una tal explicación causal tiene por su parte la cualidad de una experiencia específicamente teórica, ya que se apoya en regularidades que constata en la correlación entre determinadas convicciones morales y sus condiciones no morales.

El resultado de esta primera orientación sería, por tanto, que la posibilidad de una experiencia específicamente moral parece esfumarse desde ambas perspectivas. Si pasamos de la óptica interna, por consiguiente del propio sentido de las convicciones morales, resulta difícil ver cómo podría consentir la pretensión absoluta de los juicios morales una fundamentación en la experiencia. Si, por otra parte, pasamos a la experiencia teórica de la historia, parecemos abocados a un relativismo que derriba ya la pretensión absoluta de las convicciones morales y, en consecuencia, también la posibilidad de un proceso de aprendizaje moral. Por

supuesto, esta relación de consecuencia, por la cual el relativismo excluye una experiencia moral en tanto en cuanto contradice toda pretensión de absolutidad de la moral, muestra que, al menos, entre esta pretensión de absolutidad y la posibilidad de procesos de aprendizaje moral no puede haber la menor contradicción, sino que la experiencia moral parece implicar una determinada pretensión de absolutidad, con tal que sepamos captarla correctamente.

Para poder valorar adecuadamente nuestra problemática, parece obligado distinguir entre dos perspectivas desde las cuales se pueden discutir los problemas morales, por una parte la perspectiva de la primera y la segunda persona y, por otra, la perspectiva de la tercera persona. Ya he aludido a esta distinción al distinguir entre una óptica interna y una óptica externa. En los enunciados sobre moral formulados desde la perspectiva de la tercera persona, por lo tanto desde la óptica externa, es característico que las personas que tienen determinadas convicciones morales sean designadas como sujetos de tales enunciados, y que sus convicciones aparezcan en discurso indirecto; por ejemplo: «La persona X o el grupo X cree que tal y tal cosa es buena, porque...», y este «porqué» se refiere a la causa de su convicción. Si, por el contrario, hablo en primera persona, no ocurre forzosamente que haga referencia a mi persona, y en el enunciado no se expresa una convicción moral en discurso indirecto, sino que se expresa simplemente, por ejemplo, «tal y tal cosa es buena», y si prosigo ahora con un «porqué», este «porqué» no se refiere a la causa por la cual tengo tal convicción, sino a un motivo para el contenido de esta convicción.

Con respecto a otra persona, tengo dos posibilidades, o hablar sobre ella o con ella. Si hablo con ella,

por lo tanto en segunda persona, la perspectiva es la misma que cuando hablo en primera persona: la convicción que la otra persona expresa parece entonces una posible convicción mía; debo asumirla o rechazarla; y si entre los dos aducimos motivos en favor y en contra, se trata de motivos opuestos para un mismo contenido de la convicción. Mi intento de integrar el conocimiento de la historia de la moral en la misma reflexión moral presupone que es posible transformar la perspectiva de la tercera persona en la de la segunda persona, de forma que pueda ser vinculada con la perspectiva de la primera persona. Al final de esta lección podrá constatarse de qué modo esto es posible. Por ahora me limito a hacer las siguientes afirmaciones: el relativismo parece una posición comparativamente simple en relación a la moral sólo en tanto en cuanto nos refiramos a otros, pues entonces tenemos la alternativa de hablar también en segunda o tercera persona. Con relación a sí mismo hay mucha gente que se considera relativista, pero me parece muy difícil ser realmente un relativista moral, pues esto significaría no tener en general opiniones morales. No es posible tener opiniones morales que se puedan estimar como relativas, pues con ello estarían descalificadas como opiniones. Podemos estar muy inseguros, tanto sobre nuestras opiniones morales como sobre nuestras opiniones teóricas, pero cuando estamos seguros de que no tienen ninguna pretensión absoluta, no podemos ya mantenerlas. Un auténtico relativista moral no podría por tanto utilizar, excepto en discurso indirecto, un vocabulario moral. Sólo podría expresarse en proposiciones que indicasen sólo preferencia subjetiva, como «me gusta», y por consiguiente tampoco podría formular exigencias a los demás, que, como hemos visto, son constitutivas de las normas sociales.

Por tanto, una vez dejamos de lado el relativismo, volvemos a la perspectiva de la primera y la segunda persona y nos preguntamos desde esta perspectiva cómo podría entenderse la posibilidad y estructura de un proceso de aprendizaje moral. Ya hemos visto que aquello que hace parecer difícil la comprensión de un proceso de aprendizaje moral es la pretensión absoluta de las convicciones morales. Parece depender, por tanto, de la forma en que entendamos exactamente esta pretensión absoluta el que se dé la posibilidad de un proceso de aprendizaje moral para la concepción moral correspondiente. En la primera lección he intentado mostrar que las diferentes concepciones de la moral se distinguen mediante las diferentes propiedades que, según ellas, debe tener una norma social para estar fundada. Es naturalmente esta propiedad en la que se muestra cómo se entiende en cada concepción la pretensión absoluta de sus juicios morales.

En el contexto actual se trata en particular de las diferencias entre dos clases de semejantes concepciones. Una abarca las concepciones religiosas o tradicionalistas, para las cuales una norma social se fundamenta por la revelación divina o por la tradición. El predicado relevante es que la norma ha sido querida por Dios, es santa o algo similar, y ahí está entonces la pretensión absoluta. La otra clase abarca a aquellas concepciones para las cuales una norma social está fundada si es buena [*gut*]; ya hemos visto que este predicado puede entenderse de diferentes modos. Pero desde un cierto punto de vista estos diversos significados constituyen una clase unitaria, claramente diferente de las concepciones tradicional-religiosas. En los tres significados de «bueno» que he distinguido está contenida una condición que ha de cumplir una norma para que pueda atribuírsele este predicado, y ha de probarse

si la norma satisface esta condición. Así por ejemplo, tiene que probarse en un caso si la norma satisface la condición de ser igualmente buena para todos, en otro caso si satisface la condición de ser buena para la comunidad.

Por el contrario, los predicados «santo» y «dado por Dios» no designan condiciones o principios que deban ser controlados en virtud de las normas. Más bien, en una moral tradicionalista estos predicados corresponden sencillamente a las normas. Si bien las normas están fundadas cuando les corresponden estos predicados, estos predicados no contienen un principio de fundamentación en el sentido de un principio de enjuiciamiento [*Beurteilungsprinzip*]. Por ello, en una moral tradicionalista las normas deben ser aprendidas singularmente, no se pueden generar mediante la aplicación de un principio de enjuiciamiento. Se pueden distinguir ambos tipos de moral llamando al primero moral autoritaria y al segundo moral racional. Ahí tiene por supuesto la palabra «racional» un sentido totalmente débil. Significa meramente que se puede aducir un motivo por el cual le corresponde a la norma el predicado de fundamentación. Entonces, se puede evidenciar dentro de esta clase de concepciones morales, que son racionales en este débil sentido racional, un concepto más estrecho de «racional», a consecuencia del cual una moral semejante es racional cuando su principio de enjuiciamiento se basa en la cuestión de si es racional para todos los individuos asentir a las normas; y una moral racional en este sentido más estricto obviamente sólo es aquella en la que el predicado de fundamentación es «igualmente bueno para todos». (Para evitar posibles equívocos, recuerdo que también ésta es sólo una racionalidad en un sentido comparativamente débil; no implica que el obrar según

estas normas sea racional y el obrar contra ellos, irracional.) En las páginas que siguen me voy a limitar a considerar, en el ámbito de las concepciones morales racionales en sentido amplio, esta moral racional en un sentido más estrecho, según el cual «bueno» significa «igualmente bueno para todos».

Puedo volver ahora a la cuestión de la posibilidad de un proceso de aprendizaje moral. Como la pretensión de absolutidad contenida en una moral fundada mediante la autoridad no contiene un principio que exija el empleo de la facultad de juzgar, tampoco puede darse aquí un proceso de aprendizaje moral mediante la experiencia. En una tal concepción todas las normas tienen el mismo carácter absoluto. No constituyen un sistema en el sentido estricto del término, sino un catálogo. Son enumerables y tienen que ser aprendidas una después de otra [*additiv*]. Un orden normativo de este tipo es naturalmente característico de una sociedad estática según su óptica interna, y en este sentido no tiene ninguna perspectiva de un proceso de aprendizaje.[9]

Parece por tanto que un proceso de aprendizaje moral, en general, sólo es posible donde la pretensión absoluta de los juicios morales no tiene el sentido vinculado a la autoridad, sino que implica un principio de enjuiciamiento. Pero ésta es sólo una condición

9. En esta contraposición de ambas concepciones morales ha simplificado excesivamente la situación. Es perfectamente imaginable que en una moral fundada autoritariamente, las normas, establecidas como decretos divinos o algo similar, representen a su vez principios relativamente generales que para ser aplicados exijan la facultad del juicio. En la medida en que esto es así, tendría también una moral así un componente racional en este sentido amplio, y sería susceptible también de procesos de aprendizaje. Esto me lo ha sugerido Thomas Scanlon en Princeton.

necesaria, y no suficiente. Un caso instructivo de moral racional, pero estática, es la moral de Kant. La concepción kantiana de la moral es racional (naturalmente dejo en ello a un lado su propio uso inhabitual de la palabra «razón») en el sentido más estricto ya distinguido. Pues su imperativo categórico significa que una norma está fundada precisamente si todos pueden aceptarla. Que éste sea el criterio de Kant no lo muestra con inmediata evidencia su clásica formulación del imperativo categórico, pero subyace en última instancia y además lo hace explícito el propio Kant en su reformulación del imperativo categórico que se refiere a un «reino de los fines».[10] Pero lo que nos importa en el contexto actual es que para Kant, primero, la pretensión absoluta implícita en este principio consiste en que vale *a priori*, y en segundo lugar que, singularmente, consideró que también las normas que se generan a partir de este principio de enjuiciamiento, participan por su parte de esta aprioridad. Pero esto significa que, si bien por un lado Kant se distingue claramente de una moral tradicionalista, en tanto en cuanto no considera que las normas morales singulares sean dadas de antemano como fundadas y sean aprendidas por enumeración, sino que derivan de la aplicación de un principio de enjuiciamiento, fue de la opinión que esta aplicación era perfectamente simple y podía ser realizada de una vez por todas, con lo que resulta un catálogo de normas, cierto que no al principio sino al final. Por ello no existe para Kant ninguna necesidad de una reflexión moral y, en consecuencia, tampoco ninguna necesidad de un proceso de aprendizaje moral.

Si bien la moral kantiana parte de un principio de

10. Véase Kant, *Grundlegung zur Metaphysik der Sitten*, en Edición de la Academia, 4, p. 438.

enjuiciamiento, puede decirse que por su contenido y carácter estático está aun próxima a la moral tradicionalista. Ambas son estáticas y en ambas es característico que a las normas singulares les corresponde la misma certeza que al punto de partida, al cual deben su certeza. El precio que Kant debe pagar por esta rigidez está en la considerable represión del problema de la colisión normativa, y en la prioridad no sólo relativa sino absoluta (sin excepción alguna) que ostentan sobre todas las normas positivas (relativas a la prestación de ayuda) todas las normas negativas, o sea aquellas que prohíben dañar a los demás o limitar su libertad. Esta posición liberal de no intervención —un sistema de normas primariamente prohibitivas— no es una consecuencia necesaria del imperativo categórico en cuanto tal, sino (si se prescinde de sus conexiones sociales y políticas) una consecuencia de la necesidad de liberar no sólo al principio moral, sino a toda la moral, de toda contaminación empírica (si bien Kant, como se sabe, no triunfó en este intento, pero pudo así engañarse más fácilmente).

El ejemplo de Kant nos muestra cómo, en una moral racional, no puede concebirse la relación entre el principio y las normas concretas si es que ha de poderse dar una experiencia moral. Esta relación no debe ser entendida de tal modo que las normas concretas deriven del principio de forma cuasideductiva. Este resultado negativo puede formularse en términos positivos del siguiente modo: si ha de ser posible una experiencia moral, la norma concreta —es decir, la manera en que es juzgado moralmente un caso concreto— debe ser el resultado de dos factores, a saber, el principio de enjuiciamiento libre de presupuestos y la experiencia. Pero entonces se plantea naturalmente en seguida la cuestión siguiente: ¿cómo hemos de concebir nosotros

la relación concreta que existe entre el principio y la experiencia?

El caso más simple —en el que aquí sólo quiero entrar para después dejarlo a un lado— es el de un silogismo que consiste en una premisa mayor, que expresa una norma moral, y una o varias premisas menores, que son empíricas. La premisa mayor es una proposición normativa que indica como bueno o justo un determinado estado de cosas, por ejemplo una situación de la sociedad, mientras que en las premisas menores se afirma que ciertos medios y ciertas acciones son aptos para producir esta situación. La conclusión de un silogismo tiene entonces la forma de que —«por consiguiente»— se debe obrar así. El silogismo consiste por tanto en una premisa normativa, una o varias premisas empíricas y una conclusión normativa. Ahora bien, es característico de este primer tipo de conexión entre las normas y la experiencia que la componente de la experiencia sigue siendo meramente teórica. Todos sabemos que nuestros juicios morales concretos dependen en gran medida de nuestros conocimientos empíricos. Incluso puede decirse que en una moral racional (en sentido amplio) existe un deber moral de no juzgar moralmente un caso concreto, antes de haber conocido todos sus aspectos descriptivos y causales lo mejor posible. Y resulta claro que, en la medida en que se corrige nuestro conocimiento empírico, se corrige también nuestro juicio moral concreto. Pero también está claro que, si bien esto es así, se trataría de una forma muy débil y problemática de proceso de aprendizaje moral, porque el aumento de experiencia en este caso sería meramente un aumento de experiencia teórica: en este caso, nuestras convicciones morales están determinadas puramente por lo que es enunciado en la premisa mayor, y su propio contenido no cambiaría

con el aumento de experiencia. Sólo querría hablar de un proceso específico de aprendizaje moral cuando el componente experiencial no sólo tiene la función de hacer concretamente aplicable la norma abstracta en un mecanismo deductivo, cuando por lo tanto no sólo atañe a los medios sino cuando la experiencia tiene por resultado cambiar el mismo contenido de las convicciones morales. Esto significa que el enjuiciamiento moral en una situación concreta ya no consiste en las dos componentes independientes entre sí, una normativa y otra empírica. Y significa, por consiguiente, que la estructura de esta experienica no puede ser ya elucidada según un modelo silogístico.

¿Cómo debemos concebir la estructura de la experiencia moral en el sentido más estricto así designado? Encuentro una primera sugerencia en la obra de Hare *Freedom and Reason* (Oxford, 1963, § 3). Según Hare, el proceso de maduración moral de un individuo se presenta como un proceso que parte de normas morales muy generales que luego se concretan progresivamente. Según Kant, todas las normas morales singulares tenían un carácter muy general, como por ejemplo la norma «no se debe robar». Era este carácter general también de la regla singular el que hacía posible derivarlo del principio supremo de forma cuasideductiva y definitiva, por lo que la experiencia aparentemente no desempeñaba papel alguno. Pero según Hare tenemos que distinguir entonces entre el carácter universal y el carácter general de una norma moral. Toda norma moral, incluso la más concreta, es universal en el sentido de que, cuando en un caso concreto se dice que se debe obrar así o asá, ello implica que se debe exigir lo mismo en todo otro caso igual a éste en todos los aspectos relevantes. Pero esta universalidad no es generalidad [*Generalität*] en el sentido de una generalidad

[*Allgemeinheit*] genérica que abarque muchas especies de diversos casos. La realidad moral es tan compleja que siempre hay que formar nuevas reglas, y esto no sólo porque la realidad sea compleja meramente en sus aspectos teórico-descriptivos, sino que más bien es una complejidad específicamente moral, pues cuando se decide qué es imparcial, se trata de diversos tipos de intereses y de un conflicto normativo. En un conflicto normativo, como por ejemplo el conflicto entre el deber de ayudar y el deber de no robar, partimos siempre de normas generales, es decir de normas que Kant considera ya como las definitivas, y buscamos una norma más concreta adecuada al caso concreto.

Pero ahora se plantea la cuestión de cuál es la estructura de este proceso de 'concreción' [*Konkretisierungsprozess*]. Está claro que no puede tratarse de una deducción silogística. Si una situación cae bajo dos o más normas mutuamente excluyentes, parece obvio, para describir así el procedimiento necesario, que hay que efectuar una ponderación entre estas normas. Esta concepción, predilecta por los juristas, tiene algo de justificada pero es equívoca. Está justificada en la medida en que para los fines de la formulación del juicio moral no se puede prescindir de ninguna de las normas pertinentes, sino que hay que contar con todas. Pero la idea de ponderación [*Abwägung*] es equívoca, porque una ponderación presupone una medida común o bien un punto de referencia unitario desde el cual decidir hasta qué punto debe pesar cada norma. La referencia a la ponderación ignora el hecho de que, si verdaderamente se tratase sólo de una ponderación, sería exclusivamente el sujeto ponderador quien debería decidir subjetiva y arbitrariamente el peso de las normas, con lo que hablar de ponderación anula el criterio que en realidad es el único relevan-

te: el principio de imparcialidad o de capacidad general de consenso. Esto era ya determinante para la generación de las normas abstractas; donde estas normas están en conflicto, las normas ahora unilaterales sirven sólo de expediente, al objeto de la determinación del juicio; a donde apunta el juicio es a una nueva norma adecuada al caso concreto, que se distingue sólo por la imparcialidad, y por el sopesamiento [*Abgewogenheit*] sólo en el sentido en que la justa ponderación entre las normas debe consistir precisamente en la imparcialidad.

Juzgar imparcialmente significa preguntarse cómo se tiene que resolver el problema desde un punto de vista imparcial, que tenga igualmente en cuenta a todos. Entonces es importante aclarar cómo se debe comprender estructuralmente esta aplicación del principio de imparcialidad. Este principio no es una proposición de la cual pudiera derivarse algo; sólo designa un punto de vista desde el cual se deben juzgar los problemas morales. Este principio da solamente una directiva para la facultad del juicio y no puede servir como premisa de una deducción. Parece, por lo tanto, que Kant ha sobrevalorado la potencia de su principio moral al pretender que proporcionaba un procedimiento de decisión, que sólo necesitara aplicarse a una situación cualquiera para llegar a un resultado unívoco. Creo que el principio de imparcialidad no contiene un procedimiento de decisión, que define solamente una perspectiva desde la cual juzgar, y esto no lleva, excepto en los casos más simples, a resultados unívocos. La aplicación de este principio es insegura y remite por tanto en última instancia a la decisión autónoma del sujeto juzgante. La circunstancia de que la imparcialidad no sea un principio del que pueden deducirse mecánicamente normas concretas ha llevado a ciertos filó-

sofos a explicar que este principio formal es decididamente vacío; y cuando se dice que no existe un juicio moral sin una componente de decisión, hay personas que responden: si en última instancia subsiste un residuo de incertidumbre y decisión, ¿no sería lo mismo decidir los problemas morales arbitrariamente desde el principio? Una tal concepción equivoca el sentido que puede tener una moral racional, y esta incomprensión tiene sus raíces en la moral autoritaria bajo la cual hemos crecido todos en el primer estadio de nuestra vida; es característico de esta moral que las normas no son susceptibles de duda, así como no exigir ninguna autonomía de juicio en el individuo. De ahí nuestra tendencia —que se ha expresado claramente con Kant— a querer asimilar el principio moral racional, que es un principio de juicio, a un principio deductivo, y a afirmar que la admisión de un residuo de incertidumbre y decisionismo equivale a una decisión arbitraria. Si bien es un principio abierto, el principio de imparcialidad tiene la pretensión de absolutidad, característica de toda moral —una objetividad que no está fundada empíricamente—, e incluso las expresiones morales concretas, que son el resultado de la aplicación de este principio, tienen esta pretensión absoluta, pues si no la tuviesen, si no pretendiesen la objetividad, tampoco podrían ser inciertas.

Llego con ello a aquella característica del juicio moral que apunta a la imparcialidad, decisiva en nuestro contexto, a saber: que es corregible. Parece claro que se puede hablar de un posible proceso de aprendizaje moral sólo si nuestros juicios morales son falsables y corregibles, y si lo son desde su propia e interna perspectiva, y por lo tanto según su propio sentido. La circunstancia de que las expresiones morales concretas se fundan en un enjuiciamiento y no en una

deducción no excluye que haya aquí argumentos. Pero la forma característica que adopta un argumento contra un juicio moral no está en que este último se basa en una deducción falsa, sino en que no había tenido en cuenta un aspecto relevante. Como todo juicio moral racional en sentido estricto apunta a la imparcialidad, pero también obviamente a captar adecuadamente —es decir, completamente— el problema práctico, la admisión de que un problema moral se ha enjuiciado partidariamente —o de forma inadecuada por otro motivo— lleva a la retirada del juicio y exige una repetición del juicio tendente a la imparcialidad, una repetición que incluya los aspectos hasta ahora no considerados. Este proceso de rechazo y repetición con inclusión de nuevos aspectos puede ser repetido con tanta frecuencia como se quiera, en tanto en cuanto se pueda mostrar siempre nuevos aspectos que no habían sido tenidos en cuenta, con lo cual se muestra cómo se puede llegar a un proceso de aprendizaje moral.

Intentemos aclarar la estructura de esta experiencia moral, o al menos sus principios. Aparentemente consiste en una ampliación, en cada caso, de la experiencia teórica, en una ampliación de nuestro conocimiento descriptivo de la cosa, pero esta experiencia teórica no aparece ahora, como en el caso precedente, en calidad de premisa de una deducción, sino que sólo sirve para la ampliación del ámbito fenoménico para un nuevo juicio imparcial. El nuevo juicio es el resultado no de una deducción, sino de un nuevo acto de la facultad de juzgar tendente a la imparcialidad, y en esta medida es el resultado de una experiencia específicamente moral. La condición para que pueda darse un proceso de aprendizaje de este tipo es doble: en primer lugar, que la actitud moral se debe fundar en un principio de imparcialidad del juicio y además este principio se entien-

da como un principio del juicio y no como un principio deductivo. El principio normativo que posibilita un tal proceso de aprendizaje no puede ser escogido arbitrariamente, sino que debe ser un principio de adecuación y, al mismo tiempo —por lo que concierne a la consideración de los interesados—, debe ser un principio de imparcialidad, porque sólo una expresión normativa que pretenda la adecuación y, eventualmente, la imparcialidad puede ser invalidada con la objeción de que ha pasado por alto ciertos aspectos. En este sentido, el principio de imparcialidad tiene una estructura dinámica, que resulta autodestructiva para las opiniones normativas de que se trate en cada caso. Una experiencia o un proceso de aprendizaje sólo puede darse en general sobre la base de un principio semejante. Algo análogo vale para la experiencia teórica. También ésta se basa en un principio dinámico, autodestructivo, el principio de regularidad o conformidad a leyes. En ambos casos es la pretensión universal la que apunta hacia delante, pues tan pronto como muestra que una expresión sólo pretende ser universal, pero que en realidad es limitada o relativa, pierde su validez y remite a otra expresión en la que esta limitación queda superada. En los dos casos se trata, por tanto, de un proceso de desrelativización, un proceso que sólo es posible si previamente se da una pretensión de universalidad, pretensión que sigue vigente en el proceso. Todo proceso de aprendizaje —tanto moral como teórico— consiste en un proceso de desrelativización. Pero obviamente, resulta claro que entre ambos tipos de experiencia sólo existe una analogía formal; se trata de dos principios universales diversos, y por este motivo tiene también la experiencia moral otra estructura que la teórica: la desrelativización tiene un diferente sentido.

La prueba de fuego para una teoría de los procesos de aprendizaje moral debería ser su capacidad para integrar aquella experiencia que se puede hacer en tercera persona en relación a las condiciones causales de una convicción moral. Al principio de esta lección insistí en la diferencia subsistente entre la perspectiva de la tercera persona, por un lado, y la de la segunda y la primera persona, por otro, y todo lo que vino a continuación fue contemplado desde esta última perspectiva. Pero ésta sólo puede considerarse la perspectiva decisiva si consigue incluir la perspectiva de la tercera persona. Afirmo que es precisamente la estructura de los procesos de aprendizaje moral, tal y como ha sido descrita aquí, la que hace posible la inclusión de la perspectiva de la tercera persona en la de la primera y la segunda persona.

Ya hemos visto que el resultado de una explicación causal es la relativización de una convicción moral. Pero naturalmente queda la posibilidad de someter a una crítica tal no sólo las convicciones de los demás, sino también las propias convicciones morales. Cuando una persona cobra conciencia de que sus juicios morales son relativos a determinadas condiciones —por ejemplo, que él mismo tiene determinados intereses o de que pertenece a una determinada clase social con determinados intereses— tiene entonces la experiencia moral de que los juicios que hasta ahora ha tenido por imparciales, son partidarios, y entonces se ve obligado a asumir una nueva perspectiva, en la que se supere este partidismo.

Esta afirmación podría parecer fácilmente ingenua, pues todos sabemos que no es fácil abandonar una posición moral a la que están asociados los propios intereses, bien individuales o bien de clase. Pero se trata de una dificultad que atañe a todos los procesos

de aprendizaje moral. También cuando sólo se trata de advertir aspectos de un problema moral que antes no se han contemplado, normalmente esta dificultad no es meramente teórica, sino que se basa en el hecho de que los nuevos aspectos obligarían a uno a abandonar una concepción presuntamente imparcial, que mediante los nuevos aspectos resulta estar condicionada a los propios intereses. Y así resulta natural que los procesos de aprendizaje moral siempre tengan lugar contra resistencias subjetivas, y que sean relativamente raros. Por otra parte, no se puede negar que estos procesos de aprendizaje se dan en algunas ocasiones, y que hay argumentos morales. El tema de esta lección no es la medida en que estos argumentos son efectivos, sino sólo cuál es su estructura. Por consiguiente, lo que sería ingenuo sería pensar que es fácil aplicar explicaciones causales a las propias convicciones. Precisamente no es fácil porque conducen a una desvalorización de las propias convicciones, en tanto en cuanto hacen ver cómo las propias ideas de la justicia, etc., dependen de intereses. Pero de lo que aquí se trata no es de si es o no fácil, sino de si ello es —y cómo es— posible, y no he afirmado que suceda fácilmente, sino que, *si* alguien consigue relativizar su concepción concreta de la imparcialidad, y si se atiene firmemente al punto de vista moral racional en el sentido estricto —es decir, no abandona la idea formal de la imparcialidad— se ve forzado a dar a esta idea un nuevo contenido, en el cual se supere la relatividad que ha puesto de manifiesto su explicación causal.

Naturalmente esto vale también para una explicación causal en segunda persona. Si relativizamos la concepción moral de otro o de otra clase social mediante una explicación causal, sólo podemos probar que esta concepción es moralmente unilateral si dispo-

nemos de una nueva concepción, menos unilateral y parcial. Esto muestra que la descripción que he presentado hasta ahora no era totalmente correcta. Una explicación causal no puede, en sí y por sí, devaluar una convicción moral. Una convicción moral, que aspire a la imparcialidad, no puede ser relativizada de este modo, ni se puede probar su parcialidad mostrando simplemente que sólo podría surgir en determinadas condiciones socioeconómicas. Si así se procediese se incurriría en la llamada *genetic fallacy*, se cambiaría la demostración de la relatividad, de las condiciones de génesis de un juicio, por la demostración de la relatividad de su validez.[11] Una convicción moral que aspire a la imparcialidad sólo puede ser relativizada por una concepción a su vez moral, pero más amplia y menos partidista. Por ejemplo, la demostración de que la idea temprano-burguesa de un tráfico económico justo como derecho general ilimitado a la libertad contractual dependía de determinados intereses económicos de la clase capitalista, puede ser sólo un indicio, pero no todavía un argumento, de que esta concepción de la justicia era moralmente unilateral e injusta. Que esta concepción de la justicia sea injusta no lo puede demostrar su génesis, sino sólo sus consecuencias, y el juicio de que las consecuencias de esta concepción de la justicia son injustas presupone ya una diferente concepción de la justicia, una concepción más global. Por su parte, semejante concepción más amplia de la justicia sólo puede ser el resultado de nuevas experiencias, que hacen posible juzgar la cosa desde otras perspectivas, es decir, desde perspectivas que permitan prestar aten-

11. La falacia genética consiste, según la concepción positivista, en el intento de reducir o asimilar el valor de verdad de una proposición a las condiciones de su génesis. (*N. del t.*)

ción a aspectos que son relevantes para el juicio sobre la imparcialidad de la situación social y que habían sido pasados por alto por los representantes de la concepción precedente.

La concepción de la justicia temprano-burguesa tiene otros aspectos que no pueden ser relativizados por una concepción más amplia, por ejemplo socialista, si bien también la génesis de estos otros aspectos —como por ejemplo, los llamados derechos humanos— estuvo ligada a intereses. Naturalmente, se pueden imaginar fácilmente condiciones bajo las cuales no se formase la convicción de que un estado sin la garantía de estos derechos es injusto, o habría sido sacrificada nuevamente, como hoy sucede en tan gran medida. Pero el criterio de si es mejor conservar y ampliar o sacrificar estos derechos, consiste en si el punto de vista que, por ejemplo, recomienda su sacrificio, es el moralmente más global, es decir, si se atiene en primer lugar a la idea de imparcialidad y, en segundo lugar, si incluye las reflexiones precedentes, pero probando su unilateralidad. Como esto no es posible, debería considerarse la nueva concepción no como el resultado de un proceso de aprendizaje moral, sino como una regresión.

La reflexión sobre las condiciones causales no lleva por consiguiente en y por sí a la relativización de una convicción moral. Pero sería igualmente erróneo afirmar por este motivo que una explicación causal no pueda realizar ninguna contribución a la reflexión moral. Considerar una *genetic fallacy* a la implicación de la historia en la reflexión moral sería ya una *genetic fallacy* y equivaldría a negar la realidad de los procesos de aprendizaje moral. Estos últimos son esencialmente históricos, pero la experiencia histórica se desliza a experiencia en tercera persona e incurre ahora,

como experiencia puramente teórica, en una *genetic fallacy*, si sacrifica la orientación moral como tal, que es fundamentalmente la perspectiva de la primera persona.

Por supuesto, toda esta descripción de los procesos de aprendizaje moral parte de la presuposición dogmática de que las normas morales se fundamentan si son buenas en el sentido de «igualmente buenas para todos». He presupuesto aquí la imparcialidad como un principio vigente *a priori*. Pero no lo es; también él ha surgido históricamente. ¿Debemos decir que también ha resultado de un proceso histórico de aprendizaje? Pero debería tratarse entonces de un proceso de aprendizaje de tipo completamente diferente. En la próxima lección abordaré precisamente esta cuestión: ¿cómo se puede fundamentar una concepción moral como tal con respecto a las demás?

3. Moral y comunicación

La idea de que algo como la comunicación pudiera ser esencial para la moral y, especialmente, para la fundamentación de nuestras convicciones morales puede resultar chocante a oyentes norteamericanos. Por lo que puedo ver, no tiene precursores en las corrientes clásicas de la filosofía moral, si bien está estrechamente ligada a la tradición contractualista de la filosofía política. No sé si hay exponentes de esta concepción en la filosofía norteamericana actual. Por el contrario, hoy día en Alemania es suscrita por toda una serie de filósofos (y ello, en parte, por influencia norteamericana, de los pragmatistas y, en particular, de G. H. Mead). En mi opinión, el más importante de ellos es Jürgen

Habermas. Por ello, mi exposición de este tema tiene como base la confrontación con su punto de vista.

En primer lugar, debo decir algo sobre lo que significa aquí «comunicación», y a continuación indicaré los dos lugares de la explicación de la moralidad que he presentado en la primera lección en los que parece imponerse un aspecto comunicativo.

El que el enfoque comunicativo sea tan nuevo en la filosofía moral podría relacionarse con el hecho de que sólo en este siglo haya empezado la filosofía a abordar sus temas mediante una reflexión sobre el lenguaje, reflexión que anteriormente sólo había tenido en cuenta una dimensión del lenguaje, la dimensión semántica; pero si la función del lenguaje consiste en la comunicación, parece ser que el enfoque semántico de la moral sea una empresa incompleta; una reflexión filosófica sobre el lenguaje de la moral adecuada a la naturaleza del lenguaje debería comprenderse como una teoría de la comunicación. Habermas presentaría aproximadamente así la situación.[12] Las reglas semánticas del lenguaje deben ser integradas con reglas comunicativas o «pragmáticas», o bien deben concebirse a sí mismas como tales.

Creo que hay aquí confusiones que deben ser eliminadas si queremos dar un sentido claro a la cuestión de si la fundamentación de las normas morales tiene un carácter esencialmente comunicativo. El término «pragmático» tiene una triste historia. Como es sabido, fue introducido en el contexto de la filosofía del lenguaje por Carnap en referencia a Charles Morris, para designar una tercera dimensión de la teoría lin-

12. Véase el ensayo de Habermas, «Was heisst Universalpragmatik?», en Karl Otto Apel, ed., *Sprachpragmatik und Philosophie*, Frankfurt del Main, 1976, pp. 174-272, en especial, pp. 206 y ss.

güística, además de las dimensiones sintáctica y semántica, una dimensión que tendría que referirse al *uso* de las expresiones lingüísticas. Ahora bien, semejante definición de una disciplina como la pragmática presupone una concepción de la semántica como la del propio Carnap, a resultas de la cual una teoría semántica establece el significado de nuestras expresiones lingüísticas por medio de un metalenguaje. Pero en mi opinión esto implica una subvaloración de la semántica, es decir, una subvaloración de aquello que comprendemos cuando comprendemos el significado de una expresión lingüística. Como no resulta posible discutir adecuadamente este problema en el contexto actual, afirmaré simplemente de forma dogmática que se comprende exactamente una expresión lingüística cuando se comprende su uso. «Ahora bien —podría responder alguien— entonces tiene usted una teoría pragmática del significado.» Pero considero indiferente cómo se quiera denominar a algo siempre que podamos seguir distinguiendo lo que debe distinguirse. ¿Se sigue de la significación dada de «pragmático» que una teoría pragmática del significado tiene que ser una teoría comunicativa? Éste sería el caso si la referencia al uso de expresiones lingüísticas pudiera remitirse sólo a su uso en la comunicación. Pero resulta claro que esto sería falso, porque utilizamos el lenguaje no sólo para la comunicación, sino también para el pensamiento. Cuando utilizo una expresión lingüística al pensar, ésta tiene el mismo significado que tiene cuando la uso para comunicar mis pensamientos a otro. El modo en que Habermas emplea el término «pragmático» implica una subvaloración de la dimensión semántica análoga, si bien no idéntica, a la que tiene lugar en Carnap. Si comprendemos el término «pragmático» tal y como lo entiende Haber-

mas, es decir, de forma tal que una teoría pragmática tenga por argumento las reglas del uso comunicativo de las expresiones lingüísticas, no puede llamarse entonces «pragmática» a una teoría que trate de aquellas reglas que determinan su significado.

Pero, se podría responder, ¿no es parasitario el uso que hacemos de las expresiones lingüísticas en el pensamiento solitario con respecto a su uso comunicativo? Esto depende de qué se entienda por «parasitario». Si por ello se entiende que el lenguaje se aprende comunicativamente y que todo lo que podemos decirnos a nosotros mismos lo podemos decir también a los demás, entonces el uso solitario del lenguaje o, como Habermas lo llama «monológico», es efectivamente parasitario. Pero si por «parasitario» se entiende que el discurso interior consiste en realizar internamente una acción comunicativa, debemos rechazar este término. La cuestión de que aquí se trata no se refiere a la génesis o la importancia, sino que es una cuestión meramente descriptiva: concierne a la distinción entre dos diferentes tipos de reglas.

Una acción lingüística es una acción comunicativa cuando consiste en que un hablante o escritor da a entender algo a uno o varios oyentes (o lectores). Una acción lingüística no es comunicativa cuando no consiste en dar a entender algo a alguien. Una acción comunicativa puede ser imitada interiormente, cuando se imagina un diálogo: en este caso, conserva la distinción entre hablante y oyente, si bien en la fantasía. Por el contrario, una acción lingüística que sólo se utiliza para el pensamiento teórico o práctico no tiene el sentido de un dar a comprender, ni en la realidad ni en la fantasía. En este caso no asumo ambos papeles, no me comunico algo a mí mismo. Ya no está presente la distinción entre hablante y oyente.

Obviamente, hay oraciones que sólo pueden ser usadas en la comunicación, oraciones que se utilizan para acciones lingüísticas como agradecer, prometer o pedir. Pero otras oraciones, por ejemplo las oraciones que formulan una pretensión de verdad, también pueden ser utilizadas simplemente en el pensamiento. Para su uso no es esencial la distinción entre hablante y oyente.

Para evitar confusiones, propongo llamar reglas semánticas a las reglas que determinan aquel uso de una oración en el que es indiferente que sea o no utilizada comunicativamente, y reglas pragmáticas aquellas que hay que observar en una comunicación, más allá de las semánticas. Si alguien quiere denominar también pragmáticas a las reglas que determinan el significado, no voy a discutir acerca de palabras. Debería estar claro que la tesis de Habermas de que la comprensión de las proposiciones morales es esencialmente algo comunicativo se despuntaría con sólo llamar «comunicativas» a las reglas semánticas. Puede ser que la moralidad sea de algún modo algo esencialmente comunicativo, pero para que esto pueda mostrarse eventualmente, debe poder ser formulado antes como una cuestión abierta. Si es así, debe mostrarlo el objeto específico de la moral, y no se puede presuponer mediante la afirmación global de la supuesta incompletitud del enfoque semántico del lenguaje.

Una afirmación de este tipo tendría que aceptar también la crítica de no haber entendido la novedad de los métodos de la reflexión lingüística en filosofía. Porque no se necesitaba una gran perspicacia para reconocer que la función del lenguaje está en la comunicación; por el contrario, ésta era precisamente la concepción tradicional, que nunca ha sido puesta en duda. El mérito de la filosofía analítica consiste en

haber mostrado que el lenguaje no tiene sólo la función de posibilitar la comunicación de pensamientos, sino que el pensamiento mismo no puede concebirse fuera del ámbito del uso del lenguaje. Por su propia naturaleza, los conceptos remiten al uso de expresiones lingüísticas, y sólo si se puede mostrar que un concepto no puede ser pensado fuera del uso comunicativo de una expresión lingüística, resultaría significativo decir que el concepto es esencialmente comunicativo. Pero esto dependería de la naturaleza del concepto en cuestión.

Paso a continuación a la segunda de las cuestiones que quería aclarar antes, a saber: en qué lugar de la descripción de la moralidad que he presentado en la primera lección parece imponerse un aspecto comunicativo. Creo que en dos lugares.

El primero se encuentra en la explicación general que he dado de las convicciones morales. Según mi explicación, una convicción moral consiste en la convicción de que una norma social, es decir, una exigencia recíproca socialmente sancionada, está fundada en una determinada manera de obrar o de no obrar. El que la norma social está fundada se expresa, según cada una de las diferentes concepciones de fundamentación, en diferentes predicados: «dado por Dios», o «bueno para la comunidad» o «igualmente bueno para todos», etc. Así pues, en la medida en que la norma social es una exigencia recíproca, es algo esencialmente comunicativo. Pero sólo de eso no se sigue que también la convicción moral sea algo esencialmente comunicativo. En cuanto tengo una convicción moral o también cuando la manifiesto (puedo hacerlo en el pensamiento solitario), no comunico algo sino que juzgo o afirmo algo sobre la praxis comunicativa generalizada en la que consiste la norma como exigencia recíproca.

Tampoco parece ser nada esencialmente comunicativo la fundamentación de la convicción moral, en tanto en cuanto sólo consiste en la demostración de que el predicado de fundamentación corresponde a la norma. Obviamente, tenemos una situación comunicativa cuando dos personas discuten si corresponde o no el predicado relevante a una determinada norma. Pero los argumentos que se intercambian en una discusión semejante no son esencialmente comunicativos, porque los puede desarrollar igualmente bien una sola persona. Sería incluso sorprendente que la fundamentación de una proposición enunciativa tuviese que tener una naturaleza esencialmente comunicativa.

Sin embargo, parece que la situación es otra cuando dos personas no discuten sobre si corresponde o no a una norma el predicado relevante, sino sobre qué predicado es relevante, es decir, qué propiedad debe tener una norma con la que se fundamenta la propia subordinación a la norma y la exigencia de que los demás se sometan a ella. Aquí se plantean preguntas como las siguientes: «¿Cómo puedes tú fundamentar con respecto a mí tu exigencia de que me someta a esta norma?». La fundamentación de que aquí se trata no es la fundamentación de un enunciado, sino de una praxis común. Atañe a los motivos que las personas pueden tener y aducir recíprocamente para actuar de una determinada manera unas con respecto a otras. En este sentido, la fundamentación podría ser algo esencialmente comunicativo; podría ser que aquí «fundamentar» significase tanto como «fundamentar (justificar) algo *con respecto a alguien*». De ser esto así, se delinearía una explicación de por qué las diferentes teorías morales tradicionales no han reflexionado sobre el aspecto comunicativo: no lo han hecho no tanto porque no hayan reflexionado sobre el lenguaje o, si refle-

xionaron sobre él, porque sólo reflexionasen sobre la semántica, sino porque todas ellas sólo tenían presente una concepción moral con un predicado de fundamentación, y no vieron el problema de cómo se puede fundamentar el que un determinado predicado sea aceptado como el predicado de fundamentación relevante.

El segundo lugar en el que mi exposición de la moralidad puede haber mostrado algo esencialmente comunicativo está en relación con aquella concepción de la moralidad por mí favorecida —sin fundamentar aún esta elección—, la concepción según la cual una norma social está fundamentada cuando es igualmente buena para todos. En este caso, puede decirse quizá que el predicado de fundamentación representa un concepto esencialmente comunicativo, pues en vez de decir que una norma es igualmente buena para todos, puede decirse que todos puedan asentir racionalmente a ella. Esta idea de un consenso universal está presente en diferentes teorías morales que defienden la concepción de la imparcialidad: está apuntada en Kant, se encuentra en Rawls, y es el criterio decisivo para Habermas.

Según Habermas, el criterio de fundamentación de una norma está en que todos concuerden en asumirla en un discurso racional. Habermas entiende por «discurso» una argumentación racional entre diferentes personas. Con este término quiere indicar, por tanto, dos cosas: en primer lugar, se entiende la argumentación como una acción comunicativa, que acontece entre seres humanos [*zwischenmenschliche Handlung*]; en segundo lugar, el acuerdo al que hay que llegar no debe ser sólo un acuerdo fáctico, sino racional, basado en motivos o razones [*Gründe*].

Habermas ha inscrito esta concepción de la fundamentación de las normas en una teoría general de la

verdad que llama «teoría del consenso».[13] La tesis de esta teoría es que el criterio de fundamentación tanto de las teorías empíricas como de las normas morales debe ser en última instancia el consenso, por lo tanto, en el primer caso, el acuerdo de la comunidad científica, y en el segundo caso de la comunidad moral. Así, según la concepción de Habermas, el lenguaje tiene una naturaleza comunicativa, incluso en el ámbito que menos podía esperarse pues, como ya he dicho, las acciones lingüísticas que contienen una pretensión de verdad, no parecen ser esencialmente comunicativas.

Pero Habermas afirma que la verdad, o la fundamentación de la verdad, debe entenderse como algo esencialmente comunicativo. Naturalmente, la vinculación entre verificación o fundamentación por un lado, y acuerdo intersubjetivo por otro, es tradicional e indiscutible. Lo sorprendente de la concepción de Habermas es que, según ella, el acuerdo intersubjetivo no debe ser la consecuencia, sino el criterio de fundamentación de un enunciado.

Para los enunciados empíricos, esto me parece claramente falso. El criterio de fundamentación de las proposiciones enunciativas empíricas está contenido en las reglas de verificación que constituyen su significado. Verdad y verificación son conceptos semánticos. La aplicación de las reglas relevantes no tiene un carácter constitutivamente comunicativo. Y el hecho de que diferentes personas, al fundamentar así una proposición, lleguen a un acuerdo, se basa simplemente en que aplican las mismas reglas de fundamentación. La concepción opuesta, a resultas de la cual es el pro-

13. Véase el ensayo de Habermas «Wahrheitstheorien», en Helmut Fahrenbach, ed., *Wirklichkeit und Reflexion. Walter Schulz zum 60. Geburtstag*, Pfullingen, 1973, pp. 211-265.

pio acuerdo el criterio de fundamentación, está expuesta a la objeción de que el acuerdo fáctico no puede ser normativo, sino sólo un acuerdo cualificado.

Habermas ha tomado en consideración esta objeción,[14] pero las dos respuestas que ha dado a ella me parecen insuficientes. Naturalmente, afirma, debe tratarse de un acuerdo cualificado, de un acuerdo basado en argumentos. Pero cuando el acuerdo se basa en argumentos, en los motivos relevantes, estos motivos constituyen precisamente —según creo— el principio [*Grundlage*] de fundamentación del enunciado, y no el acuerdo. Habermas intenta obviar este O...O, afirmando que hay algo, que él denomina lógica de la argumentación, y que las reglas de la argumentación son esencialmente comunicativas. Pero nunca ha indicado en qué consisten tales reglas, y dudo que existan. Es cierto que la historia de la ciencia empírica es un proceso comunicativo, pero donde hay en este proceso desarrollos que sólo pueden ser considerados resultado de un acuerdo, se trata de desarrollos que no están determinados por reglas y no pueden ser fundados racionalmente. Se delinea así una situación con un perfil que se vuelve aún más nítido en el ámbito de la moral, a saber: hasta donde llega una justificación mediante razones [*Gründe*], permanecemos en la dimensión semántica y, tan pronto como aparece algo irreductiblemente comunicativo, no se trata ya de una cuestión de argumentación, sino de decisión, no tiene ya un carácter cognitivo, sino volitivo.

Habermas intenta determinar la cualificación del acuerdo todavía de un segundo modo. Un acuerdo racional se distingue de uno irracional —dice— en que el discurso del que es resultado, satisface las condi-

14. Cf. *ibid.*, pp. 239 y ss.

ciones de una «situación ideal de habla».[15] Estas condiciones deben ser condiciones simétricas entre los hablantes. Todos los que participan en un discurso deben tener las mismas oportunidades de hablar, nadie debe tener un privilegio especial —por ejemplo, que sus afirmaciones no pueden ser puestas en duda, etc. Aquí tenemos realmente reglas pragmáticas, que sólo en una situación comunicativa en general tienen sentido. Pero es fácil ver que estas reglas tienen exclusivamente la función de excluir las posibilidades de distorsión que pueden resultar especialmente cuando varias personas participan en un proceso de fundamentación. Esto explica por qué estas reglas no pueden ser interiorizadas, pero naturalmente también por qué no necesitan ser interiorizadas: porque no atañen a la estructura de fundamentación como tal, y sólo están ahí para evitar que factores pragmáticos extraargumentales —que sólo pueden interferir si más personas participan en el proceso de fundamentación— influyan sobre éste.

Llego así a la conclusión de que la teoría del consenso es inaceptable como teoría general de fundamentación. Pero la fundamentación de las normas constituye una situación particular —al menos en el caso especial en que el criterio de su fundamentación debe consistir en la posibilidad de que todos las aceptemos racionalmente. Mientras que en la fundamentación de una convicción teórica el acuerdo es una mera consecuencia, es indudable que en la fundamentación de una convicción moral —en cualquier caso si la fundamentación se entiende de este modo— sí entra un acuerdo en el mismo procedimiento de fundamentación. En qué medida en este caso particular la situación es diferente

15. *Ibid.*, pp. 252 y ss.

que en el caso general lo podemos aclarar considerando que, en este caso especial, podemos hablar de un acuerdo en dos ocasiones: en primer lugar, aceptamos que, si la convicción está fundada, los demás estarán de acuerdo con ella del mismo modo que con toda otra fundamentación (éste es el acuerdo que he denominado de mera consecuencia); pero, en segundo lugar, este acuerdo debe ser bien distinto de aquel acuerdo a que hace referencia la misma fundamentación. Si me pregunto a mí mismo pensando a solas: «¿Está fundada esta norma?», pregunto: «¿Puedo estar de acuerdo con la consecuencia universal de esta norma, y podría estar de acuerdo también cualquier otro?». Realizo por tanto, mentalmente y con carácter hipotético, un procedimiento universal de fundamentación, y por eso podría parecer que el procedimiento de fundamentación sería más adecuado si no fuera hipotético, sino plasmado en un discurso real con todos los implicados.

¿Lo sería realmente? El motivo por el que esto parece dudoso es, obviamente, porque aquí se tiene que preguntar también lo siguiente: el acuerdo de que se trata, ¿es sólo el acuerdo efectivo (de todos los implicados), o debe estar cualificado de un modo determinado? Para mí está claro que tiene que ser cualificado, pero de modo diferente que en el anterior. Allí sólo contaba el acuerdo que se basaba en el respeto a las reglas de fundamentación; aquí cuenta sólo aquel acuerdo que se compone de las dos fases en que se define esta concepción de la moral: cada cual ha de preguntarse a sí mismo si puede estar de acuerdo con la validez general de la norma con relación a su bien [-estar: *Wohl* (T.)], sólo valiendo este consenso como consenso racional; y, en segundo lugar, debe conceder a todos los demás el mismo derecho, pero nada más; es decir, que cuando tiene lugar un discurso real sobre

la fundamentación de una norma, los participantes no puedan hacer valer simplemente sus intereses, sino que en su discurso deben ya partir de la idea de imparcialidad; esta idea no es el resultado, sino el presupuesto de un discurso semejante. El presupuesto de un discurso así es que todos estén dispuestos a ponerse mentalmente —de forma hipotética— en el lugar de todos los demás.

Sorprendentemente, Habermas ha admitido esto, al asumir como regla de la situación ideal de habla la condición de que exista una simetría entre los participantes no sólo en lo referente a las oportunidades de hablar, sino con respecto a sus oportunidades de actuar, a sus derechos de «ordenar y de oponerse».[16] Esto significa que la comunicación que tiene presente Habermas presupone la imparcialidad y no la puede fundamentar. Se presupone entonces simplemente que la fundamentación de las normas consiste en fundarlas como igualmente buenas para todos, pero entonces se trataría de un presupuesto semántico con respecto al predicado de fundamentación, presupuesto en el que funda exclusivamente, a su vez, el hipotético discurso ideal.

Aquí se plantean dos cuestiones: primera, este presupuesto de imparcialidad, y por tanto el presupuesto de que el predicado de fundamentación relevante es «igualmente bueno para todos», ¿no debe ser fundado a su vez y, eventualmente, de forma comunicativa? Pero entonces esto no sería una fundamentación en un discurso comunicativo, que por su parte ya tendría que presuponer la situación ideal de habla. Esto atañe a mi problema relativo a la posibilidad de una fundamentación de la moralidad en general y no a una deci-

16. *Ibid.*, p. 256.

sión entre diferentes concepciones de la moralidad. Después volveré sobre este tema.

En segundo lugar, si la fundamentación de las normas concretas con el presupuesto de la imparcialidad presupone un discurso cualificado de determinada manera, ¿qué queda de la afirmación de que esta fundamentación es esencialmente comunicativa? Si el discurso moral presupone que todos los participantes tengan que estar dispuestos a ponerse hipotéticamente en la situación de los demás, ¿qué ventaja tiene entonces un semejante discurso frente a la posibilidad de que una persona pueda realizar también todo el proceso de fundamentación en su propia cabeza?

Una posible respuesta, que sugiere el mismo Habermas, es: es necesario un discurso real para informarse de primera mano sobre los intereses efectivos de todas las personas implicadas. Pero, ¿basta esto para afirmar que el procedimiento de fundamentación tiene que ser comunicativo? Voy a poner un ejemplo: en una familia se plantea un problema moral entre los hijos; interviene la madre; ésta hace exponer a cada hijo su propio interés y, a continuación decide, en calidad de observadora benévola e imparcial, cómo hay que resolver el problema. En este caso —tal como se exigía— el proceso de fundamentación moral se ha constituido sobre la base de una información de primera mano sobre los intereses en juego. Pero los actos de habla de los hijos sólo sirven de input de información. El propio procedimiento de fundamentación no ha sido comunicativo. ¿Hubo aquí algo anormal? Pues bien, opino y supongo que todos diríamos que sí hay algo anormal, a menos que se trate de niños muy pequeños, y ciertamente todos diríamos que un procedimiento semejante para la institucionalización de normas sancionadas jurídicamente no sería moralmente justificable.

Pero, ¿por qué no? ¿Por qué pensamos que las personas implicadas tienen que decidir las normas bajo las cuales han de vivir? ¿Cuál es el sentido de este «tienen» [*müssen*]? Me parece claro que no se trata de una necesidad que resulte de la regla de fundamentación, sino de un «tener que» en el sentido moral. Hasta que sólo se trataba de aproximarse a un resultado lo más justo posible, con los mejores motivos posibles, nada había que objetar al procedimiento no comunicativo de la madre. La decisión a la que ésta llegó estaría probablemente, por lo que hace al contenido, mejor fundada que si los propios hijos hubieran llegado a un acuerdo. Y sin embargo todos pensamos —supongo— que los hijos deberían haber decidido por sí mismos. ¿Por qué? Evidentemente, sobre la base de una norma moral que prohíbe la inhabilitación moral [*Entmündigung*].

Esta norma se sigue del principio moral, aquí presupuesto, de que sólo está fundamentado aquello a lo cual todos pueden asentir racionalmente. Así, supongo que ninguna persona racional, capaz de reflexionar, estaría de acuerdo en que se le impidiera decidir personalmente lo que es bueno para sí; naturalmente, puede pedir consejo a otros, pero estará dispuesta a correr el riesgo de equivocarse decidiendo qué cosa sea buena para él, antes que dejárselo imponer por otro. Supongo, por tanto, que todos queremos ser, en este sentido, autónomos, decidir personalmente por nosotros mismos. Pero de ahí se sigue entonces —por universalización— que todos queremos que sea una norma moral la prohibición de prescribir a los demás qué es bueno para ellos, aunque se pretenda conocerlo mejor. Era esto a lo que aludía cuando hablaba de la prohibición de la inhabilitación moral.

En este punto querría interponer una observación.

Hay filósofos (por lo demás, Habermas es uno de ellos) que definen de antemano la moral de tal modo que ésta se limita a normas que regulan los conflictos interhumanos de intereses. La consecuencia de una concepción semejante es que excluye por definición la posibilidad de que estén moralmente fundadas las normas que prescriben a los individuos cómo deben regir su propia vida. Pero como ya he dicho en la primera lección, no deberíamos excluir, mediante definiciones, cosas que queremos excluir moralmente. Es una realidad social que los seres humanos siempre —y, lamentablemente también hoy— tuvieron y tienen convicciones morales que atañen a la propia conducta de vida de los individuos, aun cuando ésta no ataña negativamente a ninguna otra persona. La definición amplia de la moralidad que he presentado en la primera lección admite esta posibilidad, y tiene por consecuencia que, si discutimos racionalmente con personas que tienen tales convicciones, discutimos con ellas en un sentido moral, y no semánticamente sobre el significado de la palabra «moral», lo que obviamente sería por completo absurdo.

Esta norma que deriva de un principio de la moral racional en sentido estricto y que ordena considerar a toda persona como ser autónomo que decide por sí mismo —y por tanto prohíbe dictarle qué es lo bueno para ella— tiene por consecuencia que, allí donde sea preciso decidir con respecto a una norma social —en particular, sobre una norma jurídica—, la introducción de esta norma sólo está moralmente justificada si es fruto de una deliberación colectiva en la que participen todas las personas interesadas. Puede llamársele principio de la autonomía colectiva o de la autodeterminación colectiva; pues lo que este principio prohíbe es la inhabilitación [moral] del colectivo. A resultas de

esta norma es moralmente erróneo que una persona o un grupo apelen a su propia superior sabiduría moral para decidir qué normas jurídicas deben valer, sin interpelar a las demás personas implicadas. Éste parece ser el motivo por el que los problemas morales —y, en especial, los problemas de la moral política— deben estar fundados necesariamente en el discurso de todos los implicados. Contrariamente a la opinión de Habermas, el motivo no es la naturaleza comunicativa del proceso de fundamentación moral, sino al revés: una de las reglas que resultan del proceso de fundamentación moral —que, como tal, puede realizarse también en el pensamiento solitario— prescribe que sólo están fundadas moralmente aquellas normas jurídicas que se han introducido sobre la base de un acuerdo de todas las partes implicadas.

De aquí se deduce también que el aspecto irreductiblemente comunicativo no es cognitivo, sino volitivo. Lo que exige un acto efectivo de acuerdo, de consenso colectivo, es el respeto moralmente prescrito de la autonomía de la voluntad de todas las personas implicadas. Pero este acuerdo ya no es, como en los dos casos antes considerados, un acuerdo cualificado. Ciertamente, queremos que el acuerdo sea racional, sea un acuerdo fundado en argumentos y, a ser posible, sobre argumentos morales, y, sin embargo, lo definitivo en última instancia es el acuerdo fáctico; por ello, no hay derecho a pasarlo por alto sobre la base de que no ha sido racional. Los otros dos tipos de acuerdo no eran irreductiblemente comunicativos, tenían su contrapartida monológica; por el contrario, tenemos aquí una acción irreductiblemente comunicativa y ello precisamente porque no se trata de un acto de la razón, sino de un acto de la voluntad, de la decisión colectiva. El problema de que aquí se trata no es un problema de

fundamentación, sino del problema de la participación en el poder que decide qué es lo legalmente permitido y lo que no. La idea de una ética comunicativa concebida por Habermas debe entenderse en conexión con la idea de democracia y con la exigencia de una extensión de esta idea a todos los sectores de la sociedad, con la exigencia de participación;[17] pero la ambigüedad de su teoría —que ha confundido el discurso hipotético con el real— le ha llevado a descuidar el factor volitivo y, con él, el problema del poder. Ha pasado por alto el hecho de que la argumentación moral hipotética en general no es esencialmente un discurso real, algo realmente comunicativo, y por ello —porque su uso del término «discurso» ha obviado la distinción entre esta argumentación moral comunicativa hipotética y un discurso realmente comunicativo— ha podido no tener en cuenta aquel aspecto no racional y volitivo del discurso real que exige la moral.

Creo, sin embargo, que en la concepción intuitiva del carácter comunicativo de la moral presentada por Habermas se encuentra un aspecto importante que hasta ahora no he tomado en consideración y que Habermas no ha expuesto explícitamente —ni podía hacerlo a partir de la forma en que ha planteado el problema. Ya hemos visto que, una vez presupuesta la concepción de la imparcialidad, es decir, de la universalizabilidad del consenso racional, la fundamentación de nuestras convicciones morales no es algo esencialmente comunicativo y que, por su parte, esta fundamentación conduce a una norma moral que ordena la institucionalización de procedimientos participativos de decisión, en

17. Véase, por ejemplo, su carta a Robert Spaemann, «Die Utopie des guten Herrschers», en Jürgen Habermas, *Kultur und Kritik*, Frankfurt del Main, 1973, pp. 378-388.

vez de relaciones de poder unilaterales. Afirmo que ésta es la situación si presuponemos la concepción de la imparcialidad o universalizabilidad. Pero tengo la impresión de que Habermas piensa además que, en general, sólo llegamos a la idea de imparcialidad mediante un proceso comunicativo; y, en efecto, si debe quedar algo todavía de una fundamentación esencialmente comunicativa, esto sólo puede estar aquí. Habermas no ha desarrollado su concepción en esta dirección porque ha definido la situación comunicativa en la que se discute la fundamentación de las convicciones morales mediante las condiciones de la situación ideal de habla; como estas condiciones ya implican la imparcialidad, la comunicación, así definida, no podía fundar a su vez la imparcialidad.

La cuestión de cómo se puede fundamentar una concepción de la imparcialidad pertenece naturalmente al problema más general de cómo se puede fundamentar una posición moral en general, y de cómo se debe decidir racionalmente entre las diversas concepciones morales. En la primera lección intenté mostrar que las diversas concepciones fundamentales de la moralidad se distinguen por los diversos predicados que según ellas deben convenir a una norma para que ésta pueda considerarse fundada. La fundamentación de una convicción moral en el seno de una concepción moral radica en la fundamentación de la pretensión de verdad, contenida en la convicción moral, que se expresa en una proposición en la que el predicado de fundamentación es aplicado a una norma. Por el contrario, según mi tesis, la fundamentación de la propia concepción moral no tiene ya el sentido de una fundamentación *de* (un enunciado), sino que es una fundamentación *para* el ingreso en una praxis intersubjetiva: se fundamenta que, cuando a un sistema normativo le

corresponde el predicado en cuestión, esto representa un motivo para someterse espontáneamente a este sistema, y asentir a las sanciones que lo constituyen.

¿Cuál es la configuración de una fundamentación semejante? Normalmente, se apela a una verdad superior que concierne a la propia concepción de uno mismo. Por ejemplo, se presupone la verdad superior de que somos hijos de Dios. Una vez se presupone esto, tenemos todos los motivos para asentir, en la práctica, a aquel sistema normativo dado por Dios. También puede presuponerse la verdad superior de que somos partes de un organismo. Bajo esta presuposición tenemos todos los motivos para aprobar todas aquellas normas cuya validez es buena para el bienestar de este organismo. Así pues, la fundamentación práctica de la relevancia de cada predicado de fundamentación parece tener lugar normalmente al mostrar que tenemos todos los motivos para aprobar las normas así caracterizadas si nos concebimos a nosotros mismos de una determinada manera.

Naturalmente, tampoco esta fundamentación es comunicativa en ningún sentido evidente. Es más importante constatar positivamente que una tal fundamentación es siempre ostensiblemente relativa, es decir, recurre a un presupuesto. He denominado a estos presupuestos «verdades superiores» porque se trata de proposiciones no empíricas, que no pueden ser fundadas a su vez y, por lo tanto, sólo pueden ser creídas. Los diversos predicados de fundamentación apoyados en ellos pierden por eso su fuerza de convicción tan pronto como se pone en duda una tradición determinada mediante una semejante verdad superior, o se encuentran otras tradiciones y se intenta pretender que los demás se sometan a las normas así fundadas. Naturalmente, individuos o comunidades parciales pueden

seguir creyendo igualmente en sus verdades superiores; éstas, no pueden ser ni refutadas ni fundadas. El problema se plantea sólo tan pronto como ciertas exigencias morales que derivan de verdades superiores evidentes en un ambiente particular [*provinziell evidenten*] son planteadas también a otras personas que no comparten esta fe. Hay en la historia de la fundamentación de las normas sociales un acontecimiento histórico, la Ilustración, que se caracteriza por la pérdida de fuerza persuasiva intersubjetiva de todas las verdades superiores. Este acontecimiento, del que ya había partido al comienzo de estas lecciones y que determina nuestra actual relación con la moral, tiene obviamente por consecuencia que la cuestión de la fundamentación de la moralidad en general y de la confrontación entre los diversos predicados de fundamentación —que he denominado cuestión externa de la fundamentación— resulta por vez primera patente y al mismo tiempo ineludible.

Ahora bien, si todas las verdades superiores a las que remiten los diferentes predicados de fundamentación no pueden ser ni fundadas ni refutadas, resulta ilusoria la idea de una confrontación racional entre ellas. En tanto los predicados de fundamentación remiten a verdades superiores, el resultado sólo puede ser un relativismo puro y simple. De esta situación se puede salir racionalmente sólo de dos maneras: o se renuncia por completo a la moralidad, o bien se busca un predicado de fundamentación que, sin presuponer una verdad superior, ofrezca un motivo para someterse libremente a un sistema normativo así caracterizado. Una fundamentación semejante es, asimismo, sólo relativa, pero ahora no relativa a una verdad superior, sino relativa a la alternativa de la renuncia cabal a normas sociales fundadas. El común denominador de estas dos

posibilidades son los propios intereses de los individuos, aquello que cada cual quiera con relación a su bien [-estar [*Wohl*] (T.)] y al bien de aquellos con quienes está ligado afectivamente.

Sólo en este lugar —por lo demás fundamental— parece asumir el sentido del procedimiento de fundamentación un carácter esencialmente comunicativo. Los individuos apelan mutuamente unos a otros, expresando que va en su interés recíproco el acuerdo en una praxis normativamente sancionada, según el cual sólo deben ser admitidas aquellas normas cuya validez sea igualmente buena para todos. Este proceso de conciliación [*Einigung*] sólo puede dar lugar a normas simétricas, pues su trasfondo es precisamente la alternativa de que no haya normas sociales; por ello, está excluida la posibilidad de que el punto de partida de cualquiera de los interesados consista en una posición privilegiada de poder fundada en normas.

Ahora pueden ponerse de acuerdo dos personas o grupos, A y B, en normas simétricas, a expensas de C, etc. Entonces, si no se recurre a la violencia, se plantea la misma cuestión en la perspectiva de C, etc., y resulta así un proceso de fundamentación comunicativo potencialmente universal en el que no se trata de fundamentar un enunciado, sino en el que cada uno alega ante cualquier otra persona que ésta, como él mismo, tiene motivo para someterse a una norma si él mismo se somete a ella. Este proceso de fundamentación es comunicativo porque no se fundamenta una pretensión de verdad, sino una praxis común, un común querer. La aserción de que está igualmente fundado para todos querer la validez de las normas, si son igualmente buenas para todos, es el resultado de este proceso comunicativo de fundamentaciones recíprocas del obrar.

El predicado de fundamentación «igualmente bueno para todos» se distingue, por tanto, de los demás en que para su sentido práctico de fundamentación (si un sistema de fundamentación tiene esta propiedad, se está dispuesto a quererlo) no remite a una verdad superior o, más exactamente, no tiene que remitir a una verdad superior; si, como sucede en Kant, este principio se presupone simplemente como dado a la razón [*Vernunftgegeben*] (en un sentido superior del término «Razón»), tiene el estatus de una verdad superior. Pero no tiene que ser así, porque puede ser fundado a su vez en el proceso comunicativo expuesto, en el que esta fundamentación no remite a una verdad supraempírica, sino a proposiciones empíricas que atañen a los intereses de A, B, C, etc. Esta concepción moral no se funda ya como las demás en una fe superior en la esencia del ser humano, sino sólo en suposiciones empíricas relativas a intereses mínimos fundamentales de las personas con respecto a sus relaciones recíprocas (por ejemplo, el interés que tienen en la institución de la promesa), suposiciones que presuponen su aceptación en un proceso de fundamentación intersubjetivo universal, y que por lo demás son susceptibles de revisión empírica.

He indicado anteriormente que no se pueden sopesar racionalmente entre sí las restantes concepciones morales. Por el contrario, aquella concepción resultante del proceso comunicativo presentado, que lleva a asumir como criterio orientativo el predicado de fundamentación «igualmente bueno para todos», puede ser comparada con las demás respecto a su fuerza de fundamentación: tampoco esta concepción moral carece de presupuestos, y no se puede decir que esté fundada en sí, mientras que las demás tienen sólo una fundamentación relativa; aquello que la distingue no es más

que la circunstancia de que se basa en *premisas más débiles*. La moral resultante de este proceso comunicativo puede denominarse moral mínima. Se ha criticado a esta moral que también presupone una cierta idea de la naturaleza humana, pero este presupuesto no tiene ya el carácter de una verdad superior; esta moral se limita (al menos en la intención; en los casos particulares esto es corregible) a aquellos aspectos de la autoconcepción del ser humano admitidos por todos, y esto es decisivo si se quieren fundar exigencias que han de ser planteadas a todos.

Históricamente, esta moral mínima es un desarrollo tardío, que sólo aparece una vez han perdido la credibilidad intersubjetiva todas las verdades superiores. Pero esto no la convierte en históricamente relativa en su pretensión de fundamentación. Quizá puede decirse que sólo una vez se han abandonado todas las convicciones provinciales [*provinziellen*] relativas a las verdades superiores, resulta visible la situación interhumana originaria como tal, que ha sido siempre una situación comunicativa de exigencias recíprocas y que por tanto sólo pueden ser fundadas legítimamente en el cumplimiento de pretensiones de fundamentación recíprocas.

Que ésta sea una moral mínima en el sentido de que se basa en premisas más débiles no significa que sea una moral débil en sus consecuencias. Y ello porque es universal, y por tanto exige una igual consideración de todos los individuos. Quizás esto pueda parecer obvio. ¿No se puede pensar una moral que se base en una fundamentación recíproca de unos pocos? Naturalmente, puede establecerse un sistema normativo semejante como institución social, puede funcionar, y siempre ha funcionado. Pero la cuestión es si es fundamentable y, si la fundamentabilidad significa fundamenta-

bilidad frente a todos, un sistema normativo que formule la pretensión de poder ser fundado en este sentido (es decir, sin presuponer verdades superiores) debe formular al menos la pretensión de imparcialidad. Ahora bien, esta pretensión puede existir injustificadamente, y normalmente se da así, pero éste es un problema del que se puede dar cuenta según los puntos de vista esbozados en la segunda lección. Un sistema normativo que pretenda ser justo en el sentido de imparcial es, por su misma naturaleza, inestable: está naturalmente expuesto a la acusación de no satisfacer su propia pretensión.

Pero —se podría responder—, ¿no deberíamos distinguir entre la cuestión de si un sistema normativo está fundado y la de si tenemos buenos motivos para aceptarlo? En efecto. Aquellos que gozan de las ventajas de un sistema político tienen mejores motivos para conservarlo en su injusticia que para transformarlo. Pero si bien aquellos que tienen el poder tienen buenos motivos para aprobar un sistema partidario [en el sentido de injusto: *parteiische* (T.)], no pueden hacerlo abiertamente, pues si exigen aceptarlo a otros tienen que —si no quieren recurrir a la violencia— formular al menos la pretensión de que está fundado. Pero esto significa que, si no pueden ya apelar a verdades superiores, deben pretender que ello es en el interés de todos, y de este modo no pueden evitar abrir una perspectiva que puede volverse contra ellos, al menos teóricamente, aunque en la práctica ello tenga éxito contadas veces.

(1981)

RETRACTACIONES

1. Discusión con Ursula Wolf

En su estudio *Das Problem des moralischen Sollens*[1] [El problema del deber moral], Ursula Wolf ha formulado una serie de importantes objeciones contra la concepción por mí desarrollada en las «Tres lecciones». Sólo voy a mencionar aquí aquellas que, según yo mismo pienso, exigen una revisión de mi concepción.

La primera objeción atañe al concepto general de moral desarrollado en la primera lección. Allí había afirmado que el «deber-tener que» [*sollen-müssen* (T.)] de las normas sociales —es decir, precisamente lo normativo de estas normas— sólo puede entenderse en conexión con las sanciones constitutivas de ellas. Que alguien «tenga que» hacer algo en este sentido significa que, si no lo hace, ha de esperar una sanción. Contra esta posición mía, Wolf propone distinguir entre normas y reglas sociales. Con la expresión «reglas sociales» quiere indicar aquellos imperativos generales cuya validez social no está constituida por sanciones, sino por «una crítica interesada o participante». A diferencia de los mandatos jurídicos, los mandatos morales

1. Berlín, 1984.

no deben ser concebidos —según Wolf— como «normas», sino como «reglas» (en el sentido así definido). Objeta a mi concepción que las sanciones constitutivas de los mandatos morales tienen que ser necesariamente sanciones «internas», como la censura o el desprecio moral, por lo tanto sanciones que apenas pueden ser definidas independientemente de la regla moral, y que no afectan a quien viola la regla si éste no la ha aprobado previamente. Pero entonces, el recurso a la sanción no contiene una explicación sino que introduce en un círculo [vicioso]. La reacción interna a una infracción de la regla es una «crítica interesada».

Desearía expresar de forma algo diferente lo que me parece el núcleo más preciso de esta objeción. La omisión a que se refiere está en la ausente diferenciación del concepto de sanción y, correlativamente, en la referencia global a las «normas sociales (normas éticas y jurídicas)». Las sanciones externas son constitutivas de las normas jurídicas y, por el contrario, las sanciones internas son constitutivas de las costumbres, o al menos de aquellas que consideramos fundadas y, por lo tanto, morales. Es cierto que aquellas normas que consideramos morales pueden hacerse valer también mediante sanciones externas, pero entonces no tienen la función de normas morales para aquellos en quienes están meramente vigentes.

El siguiente paso de Wolf, que una semejante sanción interna en general no tiene propiamente ya el sentido de una sanción, sino de una «crítica interesada o participante», no me parece convincente, pues considero oscuro este concepto de «crítica interesada». Tiene por objeto cancelar la distinción fundamental entre las reglas técnicas o las reglas prudenciales, por una parte, y por otra de aquellas que tienen una «validez social». La «crítica participante», que pueden ejercer unos indi-

viduos con respecto a otros en relación a las reglas del primer tipo, tiene el sentido de un consejo que los individuos siguen porque quieren por sí misma la praxis constituida por estas reglas, o bien porque quieren los fines que pueden ser alcanzados mediante la observancia de estas reglas. No tiene sentido decir que estas reglas tienen una «validez social» o que son observadas en virtud del reconocimiento intersubjetivo. La «validez social» de una regla sólo puede significar que su infracción está socialmente sancionada, y la sanción puede ser o la sanción interna del reconocimiento o desprecio, o bien una sanción externa. Por ello no parece consistente su referencia a «reglas sociales» —en el sentido por ella definido. La distinción que importa no es la distinción entre normas no sancionadas y normas sancionadas, sino entre normas sancionadas interna y externamente. Más adelante volveré a la cuestión de si el concepto de sanción interna conduce a un círculo vicioso (con lo cual, por supuesto, se anularía).

Una segunda objeción es que lo que resulta de mi formulación es un obrar conforme a la moral, y no un obrar moral, con una motivación moral. Sin embargo, yo no había intentado responder a la cuestión de la motivación moral. Se me puede objetar, no obstante, que mi planteamiento excluye cabalmente una respuesta positiva a esta cuestión. En cualquier caso, el recurso a la sanción interna permite una solución simple y natural a esta cuestión. Una acción no es sólo conforme a moral, sino moral, cuando su motivo es la sanción moral interna —y con ello me separo de la posición de Wolf. En este sentido, este segundo punto no es más que un corolario del primero.

La tercera objeción de Wolf se refiere al intento —realizado en la tercera lección— de una fundamentación comunicativa de la regla de oro: que es racional

someterse conjuntamente a un sistema normativo que sea igualmente bueno para todos. Wolf objeta que con la ayuda de un modelo contractual semejante, en primer lugar (en base a la objeción precedente) no se llega en general a una moral, sino a un sistema jurídico moralmente fundado. En segundo lugar, que de este modo no es fundamentable el principio de igualdad, pues las condiciones de partida de un pacto pueden ser desiguales. En tercer lugar, que no se ve cómo se puede llegar de este modo a un principio de universalización en sentido fuerte, a una implicación de todas las personas. En el proceso recíproco de fundamentación —tal y como yo lo presento— se presupone, en realidad, el mismo y universal reconocimiento de todos por parte de todos. Esta objeción es, por tanto, comparable a la que he formulado a Habermas, al afirmar que en su «situación ideal de habla» está ya presupuesto aquello que debería ser fundado (véase p. 134).

Por consiguiente, la parte positiva del ensayo de Wolf (a partir del capítulo 4) está dedicada a la cuestión de cómo se tiene que fundar de otro modo o, al menos, hacer plausible, una moral del reconocimiento igual y universal. De la serie de ideas en ella presentadas, sólo quiero mencionar las tres tesis que considero más importantes.

En primer lugar, Wolf intenta determinar más exactamente el alcance de la demostración. Una moral ilustrada consiste en la consideración [*Achtung*] igual y universal. El concepto de «consideración» ocupa por tanto una posición central. Esto lo considero correcto. Pero entiendo que el uso que de este término se hace en filosofía moral oscila entre dos significados que Wolf no ha distinguido con suficiente claridad.

Un significado consiste en la quintaesencia del comportamiento moral, si la moral se entiende en el senti-

do de la regla de oro. En este sentido, nos comportamos moralmente con una persona si, en términos kantianos, la reconocemos como un fin en sí, es decir, si tenemos en cuenta su bien [-estar (T.)] (benevolencia). En este sentido, la consideración es amor moral (por contraposición a afectivo) o aquello que se llama también amor al prójimo (o fraternidad). Esta relación de consideración, tal y como se expresa en la segunda formulación del imperativo categórico de Kant, es la fuente de todas las reglas morales sustantivas [*inhaltlicher*] (al menos en tanto en cuanto se refiere sólo a nuestras relaciones con otra persona; para las cuestiones relativas a la justicia esta formulación no basta): que no debemos tormentar o engañar a otro, que debemos ayudarle, etc., es una consecuencia de nuestro deber de tenerla en consideración. Esta consideración del otro como fin en sí no es sólo, empero, un principio deductivo o la simple suma del comportamiento moral determinado sustantivamente, sino que es en cierto modo su quintaesencia. Esto se aprecia de forma particularmente clara en la negación: en la falta de consideración moral, en la ofensa y la humillación. Aquello que consideramos más indignante, por ejemplo en las situaciones extremas de ilegalidad en las que se priva de todo derecho a seres humanos, como en los campos de concentración, en la tortura, etc., se encuentra el testimonio adicional simbólico de la falta de consideración: la humillación, la degradación. La consideración no es, por consiguiente, simplemente un principio, sino que también es una forma de comportamiento, una actitud.

Hay que distinguir de este significado otra acepción de la palabra que está estrechamente ligada al concepto de aprecio o estima [*Wertschätzung*] de una persona. Se puede estimar (o considerar) a una persona

con respecto de una determinada propiedad, cuando ésta ejemplifica esta propiedad de forma particularmente bien: por ejemplo, la estimamos porque es un buen profesor, músico, etc. Si, por el contrario, decimos simplemente de alguien que lo estimamos, consideramos o reconocemos, esto significa que lo consideramos como persona, como ser humano. Entendemos entonces (como explicaré más exactamente después) que este ser humano ejemplifica especialmente bien aquella o aquellas propiedades que exigimos moralmente unos de otros. Le tenemos entonces por una buena persona. Precisamente al desprecio [*Verachtung*] opuesto a esta forma de consideración se refería antes la cuestión de la sanción interna. Si uno se refiere a la sanción en términos tanto positivos como negativos, puede decirse que la consideración en este sentido —y su opuesto, el desprecio— constituyen los dos polos que conforman, en su misma polaridad, la sanción interna del comportamiento moral. Para el caso especial de la moral de la regla de oro se sigue de ahí que si la actitud exigida por esta moral es la de consideración en el primer sentido (como se ha explicado en el apartado anterior), esta actitud viene exigida por la consideración —o la amenaza de desprecio en el caso contrario— en este segundo sentido.

Me parece evidente que los dos significados de la 'consideración' son claramente distintos en el plano conceptual, aunque quizá no en el empírico. La consideración en el segundo sentido se refiere a una propiedad positiva, a una prerrogativa de las personas, mientras que en el primer sentido se refiere simplemente a su ser. Ello comporta que la consideración en el segundo sentido admita grados (en este sentido se puede considerar más o menos a una persona), pero no, en cambio, en el primer sentido. Y por lo tanto, la consi-

deración en el primer sentido no tiene ningún contrario, sino sólo un opuesto contradictorio: no hay ahí desprecio, sino sólo la ausencia de consideración, que se puede denominar desconsideración [*Missachtung* o *Rucksichtlosigkeit*].

Para diferenciar ambos significados, voy a utilizar otros dos términos en las páginas que siguen. En vez de 'consideración' [*Achtung*] en el primer sentido, voy a emplear 'respeto' [*Respekt*]. También sería posible utilizar la palabra 'atención' [*Rucksichtnahme*]; pero parece referirse un poco unilateralmente a los deberes negativos. Para indicar la 'consideración' en este segundo sentido, voy a emplear el término 'estima' [*Schätzen*] (en inglés y en francés *estimation*); para referirme a su contrario emplearé el término «desprecio» [*Verachtung*].

El segundo desarrollo positivo para mí importante del estudio de Wolf es su recurso, si bien incierto y dubitante (en el cap. 6, 2b) al intento de fundamentación psicológica de la «moralidad como condición necesaria de la identidad práctica del Yo» realizado por Andreas Wildt,[2] quien a su vez parte de las reflexiones contenidas en mi obra *Selbstbewusstsein und Selbstbestimmung*) [*Autoconsciencia y autodeterminación*] (pp. 272-278). Yo no había dado a estas reflexiones un sentido teórico-moral explícito, y agradezco a Wildt este desarrollo. Sin embargo, desde el punto de vista del contenido, Wildt se ha limitado a retomar aquel aspecto de mis consideraciones que atañe a las teorías de los roles, si bien tanto antes como ahora me sigue pareciendo decisivo su aspecto axiológico. Voy a pasar

2. Andreas Wildt, *Autonomie und Anerkennung*, Stuttgart, 1982, p. 261.

por alto los detalles específicos de la concepción de Wildt y presentar la tesis tal como resulta en relación directa a mis reflexiones de entonces.

Yo había escrito allí lo siguiente: «Evidentemente, tenemos que sentirnos como dignos de afirmación, para poder afirmarnos a nosotros mismos, y nos sentimos al menos genéticamente dignos de afirmación cuando otros nos han afirmado —nos han amado y reconocido— y también estructuralmente sólo en tanto en cuanto pensamos que los demás pueden darnos [esta] afirmación» (p. 272). Con ello, tanto antes como ahora, me parece importante notar que se habla de 'afirmar' en un doble sentido. Por autoafirmación, afirmación del propio ser, entiendo (al menos en un primer plano) la relación volitiva positiva con la propia existencia, y por tanto la disposición a seguir vivo. Aquí se trata, por tanto, simplemente de una afirmación volitiva. Puede decirse lo mismo con respecto al amor: amar a una persona significa afirmar voluntariamente su ser y su bien (estar) (y querer estar junto a ella). Por el contrario, al referirse a «digno de afirmación» se refiere uno a otro sentido de la afirmación: el de reconocer o apreciar, y ésta es una afirmación que implica un juicio de valor. La tesis es por tanto «que sólo podemos afirmar nuestra vida en el sentido del querer seguir vivos», si entendemos que es digna de afirmación, es decir digna de aprecio; ello significa que la autoafirmación en el primer sentido de «afirmar», presupone o exige la «autoafirmación en el segundo sentido, el sentido valorativo, un sentido esencialmente intersubjetivo» (p. 273).

Naturalmente esta tesis no es analítica, es sólo el fruto de una intuición y tiene que ser entendida como una hipótesis psicológica que debe ser contrastada empíricamente. También Rawls llama a una proposición

similar «ley psicológica».[3] Admito que la tesis debe ser precisada un poco mejor. Su implicación genérica tiene que ser formulada en dos tesis parciales. La primera tesis es bastante obvia: el niño pequeño sólo adquiere una relación positiva con su ser en el contexto de la experiencia del amor recibido por sus cuidadores primarios. La segunda tesis afirma que podemos amar a cualquiera sólo si lo estimamos como persona, es decir, si lo apreciamos moralmente. (Naturalmente no vale la implicación contraria.) Esta vinculación parece clara con respecto al amor y a la amistad entre adultos. El caso límite del amor en el que la consideración moral es irrelevante sería el amor simiesco [*Affenliebe*]. Obviamente, el amor hacia un bebé está desprovisto de consideraciones morales (pero naturalmente no de todas las demás formas de consideración —por ejemplo el respeto [*Respekt*]), pues el bebé no está todavía abierto a esta dimensión valorativa o moral; pero así como ésta no se aplica al niño, los adultos que tienen de sí mismos una concepción moral hacen entender a éste que, por una parte, le aman incondicionalmente, pero que por otra el amor implica la estima moral; y, por su parte, el niño querría no sólo ser amado, sino además ser tomado en serio como ser moral, es decir ser *a*) respetado y *b*) estimado moralmente. Más adelante, el joven se emancipa eventualmente (si la moral no es ya tradicionalista, es decir si se tiene conciencia de poder interrogarse autonómamente por el fundamento de las normas morales) de los contenidos concretos a los cuales estaba asociada la estima de sus cuidadores primarios y, por encima de ésta, la estima del ambiente social; consiguientemente, la con-

3. Véase John Rawls, *A Theory of Justice*, Cambridge (Mass.), 1972, § 70.

dición necesaria (no suficiente) para la autoafirmación (en el sentido simple de la voluntad de seguir vivo) no es ya el hecho de ser amado o estimado, sino la posibilidad de considerarse a uno mismo como digno de estima (y, por tanto, digno de ser estimado por aquellas personas que, según la propia concepción, tienen opiniones morales justas). La formulación correspondiente «de que uno puede considerarse a sí mismo como digno de amor» no tendría sentido, pues una semejante propiedad disposicional no existe fuera de la condición mínima de ser digno de estima. La cuestión que aquí se trata puede ser formulada más claramente en términos negativos: quien se considera moralmente digno de desprecio, es decir, quien se desprecia a sí mismo, no puede amarse a sí mismo, no tiene ninguna relación afirmativa con el propio ser.

Obviamente, estos análisis son demasiado toscos, y sólo pueden mostrar la dirección a seguir. En particular, habría que aclarar qué sucede si no se da la situación de partida socializadora favorable que habíamos presupuesto, y cómo debería entenderse la relación con sí mismos de quienes, a causa de una socialización fallida, muestran una *lack of moral sense*, una falta de sentido moral, y no están capacitados para desarrollar sentimientos de culpa.[4] Además, la realidad empírica es indudablemente más compleja, conoce diversos grados y variaciones.

Lo que me interesa ante todo es que aquí se aprecia una salida al círculo vicioso que constata Wolf en el concepto de sanción interna. La sanción moral interna —y por tanto, aquello que está en la base del sentido

4. Véase D. W. Winnicott, «Psycho-Analysis and the sense of guilt», en D. W. W., *The maturational processes and the facilitating environment*, Londres, 1965, pp. 15-28.

de obligación moral— es el desprecio moral. Este concepto sería efectivamente circular o vacío si sólo tuviese el significado de manifestar a una persona que ésta no satisface determinados estándares —llamados morales— o que actúa contra ellos, sin que su comportamiento resultase influido en lo más mínimo. En este caso, el desprecio o autodesprecio no tendría el carácter de sanción. Pero nuestra consideración del desprecio significa algo más: que la persona ha fallado moralmente y por tanto no puede ser objeto de una afirmación intersubjetiva en un sentido fundamental, que también contribuye a la constitución de la amistad y el amor. Esta conexión entre la afirmación (estima) moral y la afirmación contenida en el amor y la amistad es lo que da al desprecio el tinte afectivo de una sanción. Si se dice que la educación moral se configura mediante el amor y la privación de amor, es importante distinguir entre dos casos: si el niño no ve ninguna conexión entre la moral y el amor, para él la privación de amor es sólo un castigo externo entre otros (si bien particularmente temible). De este modo no puede constituirse ningún sentido de culpa, ninguna motivación moral y ninguna comprensión del concepto de desprecio. Esto sólo sucede si el niño ve una conexión interna entre la amoralidad y la privación de amor. Sobre esta vinculación habremos de reflexionar más después. La expresión «conexión interna», ¿significa quizá que esta conexión tiene un carácter analítico? Obviamente, se puede definir el «amor» del correspondiente modo. Pero con ello ganaríamos bien poco. La posibilidad que tenemos, en general, de hablar de sanción, presupone que aquella reacción afectivo-sentimental del ambiente que el interesado debe temer no se define exclusivamente por referencia a la moralidad. Por otra parte, el carácter interno de la sanción presupone que esta san-

ción sentimental está codefinida con respecto a él. La conexión obligada es posible mediante el asentimiento volitivo, contenido en el amor o la benevolencia. Esto no puede entenderse en términos puramente conceptuales, pero lo esencial es que existe una continuidad entre aquel amor al que pertenece analíticamente la estima moral y aquel al que ésta no pertenece todavía.

Wolf no se interesa por estas conexiones, porque entiende poder determinar el concepto de la moral independientemente de un concepto de sanción. Considera relativamente plausible la tesis de la «moralidad como condición necesaria de la identidad práctica del Yo», enunciada por Wildt, pero objeta que sobre esta base no es posible fundar el carácter específico de la universalidad del respeto moralmente exigido. Esto es correcto, e incluso es inimaginable lo contrario, porque precisamente la conexión expuesta entre la autoafirmación y la consciencia de ser moralmente digno de estima (o no digno de desprecio) concierne a la moral en general, incluso a una moral tradicional, y no sólo al caso particular de una moral universalista.

Más adelante intentaré mostrar cómo el tránsito a una moral del respeto igual y universal sólo puede tener lugar en el seno de esta conexión entre la autoafirmación y la conciencia de poder ser objeto de afirmación moral, porque el «debe» [*muss*] específicamente moral se constituye sólo en esta conexión, y por lo tanto ella constituye el terreno específicamente moral que debe estar en la base de todas las particularizaciones. Wolf cree, por el contrario, que en este punto es precisa una concepción totalmente nueva. A este tercer desarrollo de sus análisis positivos está dedicado el último capítulo de su estudio. Se trata ciertamente del más rico en ideas de toda la investigación, pero en el contexto de mis «Retractaciones» no puedo adentrarme

en las particularidades de esta empresa, sino sólo considerar la dirección fundamental que toma con este último paso. La tesis es que el problema de la moral debe ser planteado en el contexto del problema de la «vida buena», es decir, en el contexto del problema de qué quiero en definitiva.

Se trata esencialmente de un retorno a la concepción platónico-aristotélica, según la cual el problema del deber moral (καλόν) debe ser reconducido o transformado en el problema de los propios intereses bien entendidos, de aquello que es verdaderamente querido (ἀγαθόν, βουλητὸν ἀληθές). En los últimos tiempos, esta vuelta ha sido también efectuada por diversos autores,[5] y quizá de forma más resuelta por Philippa Foot con su concepción de la «moral como sistema de imperativos hipotéticos».[6] Las reglas morales son reducidas a reglas prudenciales. Obviamente, para Wolf este desarrollo es una consecuencia implícita de su punto de vista, según el cual el concepto de la moral debe poder determinarse sin tener que recurrir a un concepto de sanción.

Del sentido y límites de una concepción semejante puede obtenerse simplemente una idea si uno se pregunta qué significado adquiere entonces la «obligación» [*muss*] de las normas morales. Philippa Foot da inequívocamente la respuesta exacta a este interrogante: el sentido de la obligación es ahora un imperativo hipotético; así entendidas, las normas morales son normas racionales: que yo «tengo» que actuar de cierto modo tiene ahora el sentido de que sólo si actúo así obtengo

5. Véase también mi ensayo «Ética antigua y moderna», contenido en el presente volumen, pp. 39-65.

6. Philippa Foot, «Morality as a system of hypothetical imperatives», en *Philosophical Review*, 81 (1972); también en Ph. F., *Virtues and Vices*, Oxford, 1978, pp. 157-173.

aquello que en definitiva quiero, o puedo ser tal y como en definitiva quiero ser. Todas las connotaciones intersubjetivas de la obligación de las normas sociales deben estar, por tanto, estrictamente separadas de esta [idea de] obligación: caduca así toda referencia a una obligación moral y a la culpa. Si no se obra como se tiene que obrar, aparece ahora una crítica que no contiene reprobación, censura o desprecio, sino compasión e ilustración [*Aufklärung*], como en Sócrates. Quien obra mal en este sentido, no resulta culpable, sino que es memo.

Con ello no pretendo desacreditar esta concepción. Al contrario, no considero impensable que hayamos llegado hoy nosotros (como hicieron Platón y Aristóteles) a la concepción de que la moral en sentido recibido remite a razones que no están ya a nuestra disposición, y que ahora sólo es fundamentable un concepto semejante del propio interés bien entendido. Pero lo que en cualquier caso deberíamos evitar es que connotaciones de concepto, que por su sentido pertenecen a la obligación de las normas sociales (sancionadas), sean extendidas de forma inadvertida e indebida a la obligación definida en términos de prudencia. Philippa Foot es aquí muy clara (y menos ambigua que Wolf), pero reivindica aún el sustantivo «moral». Naturalmente, es posible dar una nueva definición a esta palabra. Sobre ello me limito a decir lo siguiente: para mí sería mejor —la decisión que tendríamos que tomar aquí sería más clara— si, en vez de hablar de una concepción prudencial *de la* moral, se dijese que se intenta fundar prudencialmente un corpus de reglas de acción que hasta ahora han sido entendidas como reglas morales. Al menos esto debería estar claro: se trata de una nueva propuesta —y, si se quiere conservar para ello la palabra «moral», de una nueva propuesta relativa a la com-

prensión de esta palabra—, mientras que en las «Tres lecciones» yo estaba interesado por el problema descriptivo de cómo llegar a un concepto que se adaptase también a lo que a grandes rasgos se ha entendido hasta ahora por moral, y también a aquello que indagan los etnólogos cuando investigan la moral de un grupo; por ello, allí planteaba el problema de una nueva fundamentación de la moral en el horizonte de este concepto presupuesto y dado de antemano.

Si las reflexiones que he ofrecido a propósito de la «tesis Wildt» son más o menos exactas, este concepto se refiere a un hecho psicológico preliminarmente dado en nuestra socialización, sobre el cual se puede indagar si es en general posible o deseable prescindir, bien posteriormente en individuos que ya han sido socializados (pero ¿cómo?), bien en el futuro con otra socialización (¿cuál?). La concepción prudencial parte del presupuesto teórico de la *lack of moral sense* antes citada. Pero de hecho está sostenida por filósofas de las que yo no diría que adolecen de *lack of moral sense*, y que precisamente por eso llegan a una concepción de aquello que en definitiva quieren, que sustancialmente coincide con una parte esencial de aquello que se conforma al *moral sense*. Como intentaré mostrar después, la nueva fundamentación de la moral, tal como resulta obligada en los tiempos modernos, puede moverse aun perfectamente en el marco del concepto moral tradicional (es decir, de la obligación sancionada mediante el amor y la estima). Dudo que pueda prescindirse de los conceptos de «obligación» y «culpa», y creo que una conciencia llamada autónoma sólo puede desarrollarse sobre la base de la conciencia heterónoma infantil.

También puede decirse de la obligación moral que ésta se encuentra por su parte en el contexto de la vida buena, es decir en relación con aquello que quiero en

definitiva. De hecho, si la conciencia de ser digno de afirmación como ser humano —es decir, ser moralmente digno de afirmación intersubjetiva— es la condición para poder afirmarse a sí mismo, esto significa hacer de la moralidad un componente esencial del modo en que se quiere vivir. Puede también decirse: para poder vivir como quiero vivir, *tengo que* (y aquí la expresión tiene su sentido prudencial) orientarme según la obligación de orden moral. Pero con esta inserción en el ámbito de la obligación prudencial, la obligación moral no pierde en modo alguno su sentido propio e irreductible, que se basa en una sanción social (por supuesto particular).

2. Consecuencias para el concepto de la moral

Después de haber guiado en la sección anterior la sucesión de mis reflexiones por la exigencia de la discusión con Ursula Wolf, de lo cual resultaron una serie de acentos nuevos, sin una precisa conexión sistemática, voy a considerar ahora sistemáticamente las modificaciones resultantes para la concepción expuesta en las «Tres lecciones».

En ellas he defendido la concepción según la cual si decimos simplemente de una acción que es buena (o mala), queremos decir que corresponde a una norma (o la contradice), que es considerada buena. Afirmaba allí, por tanto, que el término «bueno», usado en sentido absoluto (como predicado simple), es un predicado que se refiere principalmente a normas; y de ello se seguía que esta palabra, así utilizada, no tiene en el lenguaje una significación unívoca. Sólo en una nota (véase p. 80) he observado que el uso absoluto de este término podría entenderse también en el sentido de

que una acción es buena si es la acción de una persona buena. Por tanto, el uso absoluto de esta palabra en referencia a acciones se basaría en este uso atributivo, referido a la denominación de la clase «ser humano». De las reflexiones previas parece imponerse la conclusión de que este uso atributivo subyace al uso moral de la palabra «bueno» porque, si podemos entender la sanción específicamente moral exclusivamente en el sentido de que cualquier persona es estimada como ser humano —y «estimar» significa siempre «considerar bueno»—, resulta inevitable este recurso al concepto «bueno como ser humano».

En las «Tres lecciones» había excluido este camino porque sólo tenía presente la significación aristotélica. Por significación aristotélica entiendo no la concepción específica de Aristóteles, sino toda concepción según la cual la moral debe estar fundada en un concepto de la esencia, función o destino del ser humano. Una concepción semejante, según la cual el deber moral se funda en una determinada tesis relativa a la esencia del ser humano, debe apelar siempre a una «verdad superior» (véase p. 141), y por lo tanto a una premisa no susceptible de fundamentación empírica, en este caso no una premisa avalada por la tradición, sino metafísica (es decir, con la pretensión de ser evidente *a priori*). No puedo tomarla en consideración con vistas a la determinación general del concepto de la moral que me propongo, porque las concepciones tradicionales de la moral no pudieron subsumirse previamente en él.

Ahora puede verse claramente que, dado el modo en que he utilizado anteriormente la expresión «bueno como ser humano» (véase p. 152), el anexo «como ser humano» sólo puede tener en realidad un sentido negativo. Apreciar a alguien como ser humano significa estimarlo no por tal aspecto o actividad suya particular

(como comerciante, profesor o dirigente), sino apreciarlo sin más. Quizá pueda parecer extraño, pero es incontestable que podemos estimar a una persona no sólo en este u otro respecto, en este o aquel rol, sino además pura y simplemente como persona. Ya hemos visto que esta estima fundamental está también en la base del amor y la amistad, y el amor y la amistad se refieren precisamente a la persona en cuanto tal y no en cuanto dotada de una cierta propiedad.

La seguridad intuitiva de que existe algo como la estima pura y simple de una persona, no nos dispensa por supuesto de aclarar qué es lo que estimamos en la persona, cuando la estimamos como tal. Naturalmente sería tentador para mí responder del siguiente modo a esta pregunta: si se considera a una persona buena sin más, esto significa (y con ello se define) que «ejemplifica especialmente bien aquella o aquellas propiedades que exigimos moralmente unos de otros» (véase p. 152). Es tentadora esta respuesta, porque entonces debería sacrificar relativamente poco de la concepción desarrollada en las «Tres lecciones». Pues entonces la definición citada puede ser ampliada sin más así: que una persona se adecúa a nuestras exigencias morales significa que sigue aquellas normas que consideramos fundadas. Esta definición es lo suficientemente general para poder admitir los diversos predicados de fundamentación, según la concepción desarrollada en las «Tres lecciones». Ofrece una explicación natural de la posibilidad de que una *persona* calificada de 'buena' (digna de estima) en *toda* concepción moral, incluso en aquellas en las que el *predicado de fundamentación de las normas* no es la palabra «bueno» en uno u otro de los significados citados en las «Tres lecciones»; y realmente éste parece ser el caso. Por ejemplo, si un sistema normativo se considera fundado si es dado por

Dios, son considerados buenos todos aquellos que se adecúan voluntariamente a él (más exactamente; a su sanción interna). Y la interpretación aristotélica (que no he tenido en cuenta en las «Tres lecciones») podría incluirse a su vez en esta concepción general, en calidad de un caso particular entre otros, con un predicado de fundamentación propio.

Naturalmente, esta concepción resulta tentadora para mí por cuanto me permite conservar al menos la *determinación conceptual* de la moral propuesta en las «Tres lecciones». Pues el modo de empleo actual de la palabra «bueno», para mí absolutamente nuevo, que se refiere al correlato de la estima de una persona en cuanto tal, no contendría según la explicación dada ningún concepto de «moral» en conflicto con el presentado en las «Tres lecciones», sino que se apoyaría en este último.

No obstante, en una reflexión ulterior esta concepción no puede satisfacer realmente. Lo que desde un principio me causaba malestar en ella era el hecho de que su identificación de «bueno» (digno de estima) sin más con «moralmente bueno», es decir, con «bueno en la observancia de las normas que se consideran fundadas», es en rigor una proposición sintética, que por tanto aun estaría por fundamentar. Pues con seguridad no es una noción sencillamente analítica la de que estimamos a una persona en sí misma en cuanto ésta observa aquellas normas que consideramos fundadas. El núcleo de este problema es que la referencia a la estima de una persona «como tal» («sin más», «en sí misma») sigue sin estar clara. No es lícito resolver la cuestión mediante un golpe de fuerza definitorio, decretando que tiene el significado de estimar moralmente a la persona; antes bien, debemos afrontar el fenómeno que se nos presenta, aunque todavía de forma vaga,

cuando hablamos, como acabo de hacer, de una certeza intuitiva como la que nos viene proporcionada por este fenómeno.

Por último, lo que me ha revelado claramente la insuficiencia de esta concepción ha sido la forma en que se refiere a la fundamentación de las normas, que se sitúa por debajo del estadio al que ya había llegado (si bien de forma no totalmente clara) en las «Tres lecciones». Había llegado allí al resultado de que apenas tiene sentido decir, con respecto de las normas, que están fundadas. No me refiero sólo al argumento aducido contra Habermas, a saber, que no tiene sentido afirmar simplemente, de una norma, que está fundada; sino también a que lo que he dicho acerca de los predicados de fundamentación guarda relación con la idea errónea de que una norma está fundada si se muestra que le corresponde el predicado de fundamentación de que se trate en cada caso. Esto es permisible como una formulación abreviada y provisional, pues ya se había mostrado que aquello que con rigor está fundado no es nunca la norma o una propiedad de la norma, sino la adopción de la praxis correspondiente: que tenemos un motivo para someternos voluntariamente a las sanciones que otorgan validez a la norma; y entretanto se ha explicado que, cuando se trata de moralidad y no de legalidad, esta sanción es la sanción interna del desprecio, y que someterse voluntariamente significa considerarse a uno mismo (y a los demás) digno de desprecio cuando se transgrede la norma o, por el contrario, como digno de estima cuando se observa su cumplimiento. Ahora bien, si el motivo [*Grund*] consiste precisamente en el hecho de que se tenía un motivo para adoptar esta praxis referida a la estima, la estima de una persona en cuanto tal no puede ser definida por el hecho de que ésta realiza

esta praxis fundada, como habíamos presupuesto en la concepción antes desarrollada.

Esto puede parecer tan complicado que se tienda a desechar la totalidad de la argumentación por su carácter artificioso. Sin embargo, estamos cerca de una sencilla solución. Recordaré un aspecto de mi explicación de lo que significa fundar la adopción de una praxis normativa, cuya significación no me resultaba aun tan clara: la idea de que, en rigor, la fundamentación se compone de dos fases (véanse pp. 96 y ss.). El individuo pregunta por qué debe asentir voluntariamente a la sanción correspondiente a un sistema normativo. La respuesta es que porque a estas normas les corresponde una propiedad destacada [*ausgezeichnete*] (precisamente aquella que se expresa en el predicado de fundamentación). Pero esta respuesta sólo tiene una fuerza motivadora y fundamentadora en asociación a una determinada concepción de sí mismo del individuo (véase p. 97). Por ejemplo, está fundado, justificado, someterse al sistema normativo que tiene la propiedad de ser mandado por Dios, si uno se considera hijo de Dios. Esta propiedad —ser hijo de Dios— no es una propiedad cualquiera, sino aquella que se considera constitutiva de la propia naturaleza.

Este aspecto, que ya se impuso por tanto en la problemática de la fundamentación, parece sugerir ahora una respuesta simple a la cuestión planteada, la de qué significa estimar a una persona «sin más». Toda concepción moral que se expresa en cada caso con un determinado predicado de fundamentación parece fundarse en una concepción de la característica esencial de la identidad de la persona (¡pero no al revés!). ¿Significa esto entonces que estamos obligados a volver a la vieja interpretación aristotélica? Mi respuesta es que no. El significado que atribuimos aquí a «esen-

cial» no deriva de un predicado determinado, no es, por ejemplo, la esencia del «ser humano» (o de la «persona»), sino que se trata de la propiedad que en cada caso «nosotros» (el grupo con una determinada concepción moral) suponemos constituye la propiedad decisiva de nuestro ser o de nuestra concepción de nosotros mismos (por ejemplo, ser hijo de Dios o miembro de este pueblo, etc.). Si se me pregunta ahora cómo puedo probar que no podemos dejar de atribuirnos recíprocamente siempre una propiedad esencial como ésta, sólo diré lo siguiente: parece tratarse ante todo de un hecho empírico, pues esta atribución recíproca es siempre el presupuesto para que podamos estimarnos recíprocamente y estimarnos a nosotros mismos «sin más». Sólo más adelante podré dar una respuesta más profunda a esta cuestión.

Ahora me veo obligado, por tanto, a ofrecer una definición completamente nueva del concepto de moral. Ya no puedo decir que «bueno como ser humano» debe ser definido mediante un concepto de «moralmente bueno» ya presentado antes en otra parte, sino que, al contrario, «moralmente bueno» es una expresión que no se puede definir de otro modo que «bueno con respecto a la propiedad considerada intersubjetivamente esencial». Obviamente, la propiedad presupuesta como esencial es tal que permite una serie de grados según el criterio bueno-malo, pues de otro modo no podríamos estimar a alguien con respecto a esta propiedad.

¿Cómo se introduce este empleo de «bueno» entre las diversas significaciones de «bueno»? Desde el punto de vista de la formulación actual del problema, obviamente la breve referencia a los diversos significados de «bueno» que he presentado en las «Tres lecciones» (véanse pp. 76 y ss.) resulta bastante desafortunada.

A propósito del uso atributivo de «bueno», además del uso estético había unido allí dos significados bajo el título de «idoneidad de algo para el desempeño de su función característica», ofreciendo como ejemplos «un buen cuchillo» y un «buen futbolista» (p. 77). Pero ahora hay que distinguir aquí (como ha hecho G. H. von Wright) entre una función referida a un fin —como la del cuchillo— y una función en la cual no se puede hablar de un fin, como en un buen futbolista, un buen bailarín, una buena madre, y similares. Lo erróneo de aquella formulación, desde mi perspectiva actual, fue que en el desarrollo perdí de vista al futbolista y me orienté sólo hacia el cuchillo. Pero el significado de «bueno» (como ser humano) que resulta decisivo ahora pertenece justamente a la clase de aquellos casos en los que algo cumple bien una función, pero una función no relacionada a fines. Esta significación está muy próxima al uso estético, al que antes consideraba que era irrelevante en el contexto de mi problemática.

En vez de éste, consideré entonces conjuntamente el uso atributivo ligado a fines con el uso hipotético y el uso en el sentido de «oportuno», porque los tres usos son relativos a fines, diciendo luego que todos los usos que no son morales ni estéticos son, en este sentido, modos de uso relativos. Es característico de estas formas de uso relativo —como mostré allí— que estas proposiciones valorativas son traducibles a proposiciones de deber, en las que «debe» significa lo mismo que «si no eliges esto, te comportas de forma irracional». Es un «deber racional». Por el contrario, aquella significación atributiva de «bueno» referida a una función no ligada a fines no es relativa en este sentido, al igual que la estética. Por eso no parece ser aquí plenamente significativo hablar de racionalidad; el

correspondiente deber no se refiere a una norma racional.

Siempre que una proposición de deber es traducible en una norma racional, no desempeña la estima un papel constitutivo; en este caso es estimado, considerado bueno, precisamente aquello que es bueno según criterios racionales. Por el contrario, aquel «debe» que podemos usar cuando se trata de roles y de otras funciones no relativas a fines remite, al menos parcialmente, a una estima que tiene una significación valorativa de carácter constitutivo. Quien baila bien, baila como se debe bailar, donde «debe» significa: de la forma apreciada en un bailarín. Esta estima no es puramente subjetiva, se refiere a estándares en relación a los cuales se puede argumentar y fundamentar, pero que en sí mismos son en último término subjetivos. (En el fútbol y similares tenemos por supuesto la medida objetiva del triunfo, y en los casos de roles también concurren en los aspectos referidos a fines; pero en todos los casos también desempeñan un papel los estándares subjetivos.)

No obstante, en todos estos casos, la estima no es una sanción. La estima es aquí siempre una «sanción interna» en el sentido restringido de que se refiere analíticamente a la función correspondiente (y por ello no es una sanción en el sentido estricto del término). Si no apreciamos a una persona como bailarín, esto significa exclusivamente que afirmamos satisface mal los estándares obligatorios para el desempeño de esta propiedad o función. Frente a esto, ya hemos visto que el rechazo de la estima moral va más allá de la pura constatación que la persona no se adecúa a determinados estándares, y por tanto es una auténtica sanción (véanse pp. 156 y ss.). Obviamente, esto está estrechamente relacionado con el hecho de que la función a que se refiere

la estima moral no es una función discrecional que alguien puede satisfacer si quiere, sino que se presupone relacionada a nuestra propiedad esencial. Y por ello aquí no decimos, como en los demás casos: tienes que obrar necesariamente así, si quieres ser así y así; sino: tienes que obrar así, y basta, sin condiciones. Si alguien falla en esta función central, no es condenado en un determinado sentido, sino que es condenado *él*. Esta estima concierne al sentimiento que la persona tiene de su propio valor.

Este uso de «bueno», ¿es atributivo o absoluto? Puede ser ambas cosas, y ahora podemos ver por qué y en qué sentido. Como habíamos visto, la estima moral concierne a la «misma» persona, a la persona «sin más». Naturalmente no pretendemos dar a entender con ello la convicción de que bajo todas las propiedades existe algún núcleo sustancial de carácter metafísico. Al igual que todo tipo de estima, esta estima puede referirse sólo a una propiedad, a un aspecto. En este sentido este uso de «bueno» es atributivo. La persona es buena —por ejemplo, en el modo en que satisface la propiedad de ser hijo de Dios. Pero si se considera esta propiedad como la propiedad esencial, esto significa que la persona es buena, y punto (en el mismo sentido en que antes pude decir: tienes que ser así, y punto). En contraste con lo que consideraba en las «Tres lecciones» un uso absoluto de la palabra «bueno» (y que en realidad, como naturalmente admito, era un uso relativo: «bueno *para* todos», «bueno *para* el grupo»), tenemos aquí un uso absoluto de la palabra, que deriva precisamente del hecho de que se presupone que una determinada propiedad es una propiedad esencial para nosotros.

Quiero ahora recapitular estas reflexiones sobre el concepto de la moral en cuanto tal a fin de poder pasar

a la cuestión de cómo tengo que entender ahora de nuevo aquella concepción de la moral que subsiste cuando —como he formulado en las «Tres lecciones»— no pueden ser ya aceptadas «verdades superiores». ¿Cómo están ligados entre sí los conceptos, esenciales en toda moral, de fundamentación, norma moral, «bien moral», sanción interna y estima de una persona como tal? Voy a intentar exponer ahora la conexión de forma tal que se sitúe en el centro el problema de la fundamentación, para mí tan decisivo ahora como antes.

El individuo en sociedad se ve confrontado a normas, normas que tiene que observar, donde la expresión «tener que» significa precisamente que tales normas van acompañadas de aquella sanción que consiste en la estima-desestima de la persona en sí misma. Las normas sancionadas de este modo son (para la sociedad) normas morales. Un individuo, que es insensible a esta sanción —que está en la base de la posibilidad del amor y la amistad— (*lack of moral sense*), no puede entender este sentido del «tener que» [observar la norma]. Para un individuo semejante, la palabra «moral» es una palabra vacía. En caso contrario, exigirá (o podrá exigir) una fundamentación, que justifique por qué él tiene que limitar su propia libertad de acuerdo con estas normas, para ser objeto de asentimiento [o afirmación] por los demás (utilizo aquí la expresión indeterminada «asentimiento» para abarcar todo el complejo de la estima y de la posibilidad de amor) [*Geschätztwerdens und Geliebtwerdenkönnens*].

La respuesta es más o menos ésta: en primer lugar, estas normas tienen una propiedad determinada (la propiedad expresada mediante el predicado de fundamentación). Que tenemos esta propiedad puede ser constatado empírica o cuasiempíricamente (el «cuasi-

empíricamente» se refiere a una propiedad como «dado por Dios»; aquí se cre que, por ejemplo, una tal revelación ha tenido realmente lugar en la historia). En segundo lugar, «todos» tenemos la propiedad esencial. En tercer lugar, esta propiedad o función sólo puede ser satisfecha (es decir, sólo se es moralmente bueno), si se observan las reglas correspondientes a aquella propiedad normativa. Este paso es analítico, resulta de la naturaleza de esta propiedad. En cuarto lugar, por consiguiente, se puede estimar sin más a alguien, sólo si observa aquellas normas.

Si, como he subrayado antes, un individuo pone en duda la moral, es decir, la sanción moral en general, en el sentido de que no tiene sensibilidad para esta sanción, no es posible discutir con él. Se ha puesto al margen de toda situación de afirmación recíproca de una forma tan fundamental, que ha hecho por cuenta propia exactamente aquello que es el sentido de la sanción moral (que por supuesto en este caso no se concibe como sanción). En mi opinión, este caso es patológico. Esta denominación no pretende ser de descrédito; por el contrario, muestra el hecho de que toda moral se apoya en una base empírica, en la presuposición empírica de una determinada constitución o estructura sociopsicológica. A este nivel tan bajo [*untersten Ebene*] no hay nada que fundar. El problema de la fundamentación se plantea siempre ya bajo el presupuesto de que una persona está vinculada al amor, a la amistad y a la estima que ambos suponen, y también de que sea capaz de estimar sin más a las demás personas. El problema de la fundamentación atañe, por tanto, sólo a la validez de las cuatro fases de la fundamentación antes citadas.

Obviamente, la segunda de estas fases es la crítica. (Eventualmente también la primera o parte de la pri-

mera, a saber, la que antes he llamado cuasiempírica; pero de ello no se deriva ninguna dificultad esencial adicional con respecto a la dificultad de la segunda fase.) ¿Por qué —puede preguntarse— es precisamente *ésta* la propiedad esencial para nosotros? Si la pertinencia de esta propiedad esencial tiene el carácter de una verdad superior, este problema en última instancia es irresoluble. Una moral que se base en una verdad superior está por lo tanto basada sólo en el sentido muy limitado de fundarse sobre un estándar aceptado por aquella que para nosotros es la propiedad esencial. [*Pero*] este estándar no puede ser a su vez fundado.[7]

3. La moral del respeto recíproco

La autoatribución de una propiedad esencial, *¿tiene que* tener el carácter de una verdad superior? ¿Qué sucede si el problema de la fundamentación llega hasta el punto de poner a su vez en cuestión la verdad superior presupuesta en una tradición? He sostenido en las «Tres lecciones» que aquello a lo que todavía se puede recurrir en este caso son los intereses empíricos

7. Ni en su crítica de mi posición ni en su propia argumentación ha tenido en cuenta U. Wolf la diferencia entre el concepto general de la moral y de la situación moral que resulta cuando no se cree en ninguna verdad superior. Así, ha ignorado también el problema del mejor fundamento de la moral de la consideración con respecto a otras concepciones morales. Y, de hecho, la fundamentación débil, «estética», que da a la moral de la consideración tiene en definitiva por resultado que ésta no puede pretender estar mejor fundada que otras concepciones morales. Cada cual tiene la moral (si puede hablarse de una moral) que corresponde a su concepción de la vida buena.

de los individuos, en particular el interés de cada cual en un sistema normativo bueno para uno. Parecía resultar de ahí un nexo de fundamentación recíproca: que para cada cual es bueno someterse voluntariamente a un sistema normativo que es bueno en igual medida para todos.

Contra esta concepción son válidas las citadas objeciones de Ursula Wolf: en primer lugar, que lo fundado así es un sistema jurídico, no una moral; en segundo lugar, que esta fundamentación presupone ya un punto de vista de la igualdad que no está dado naturalmente de antemano, sino que debería ser a su vez fundado moralmente.

La debilidad sistemática inmanente de aquella argumentación mía radica de nuevo en una falta de claridad del concepto de fundamentación. En mis ensayos anteriores (a las «Tres lecciones»), cuando no había aclarado aún el sentido práctico-subjetivo del concepto de fundamentación, había pensado, en primer lugar, que para un individuo está fundada la aceptación espontánea de un sistema normativo si éste es bueno para él y, en segundo lugar, que la norma estaba fundada si era buena para todos. En el primer paso está presupuesto el significativo concepto de fundamentación práctico-subjetiva, mientras que el segundo contiene la idea inversa de una fundamentación *de* la norma. Puede encontrarse aún un residuo de esta concepción en la primera de las «Tres lecciones» (véanse pp. 90 y ss.). En la tercera lección he evitado, consecuentemente, esta forma de hablar, sustituyéndola por la idea de una fundamentación recíproca («comunicativa») —en la que la fundamentación debería entenderse en sentido práctico-subjetivo, pero de forma que cada cual hiciese presente a los demás que es racional para él y para los demás someterse al sistema norma-

tivo, si los demás se someten también. Ahora bien, cuando retomé esta argumentación (véanse pp. 144 y ss.) afirmando que esta fundamentación debía ser entendida como una «susceptibilidad de fundamentación (de la norma) con respecto a todos», incurrí en una confusión conceptual. En realidad, en aquel proceso intersubjetivo de fundamentación puede decirse sólo que está fundado *aditivamente* para cada individuo, en cada caso, asentir al sistema normativo en las circunstancias dadas. Por el contrario, en la «fundamentabilidad con respecto a todos» resuena de nuevo el eco de la fundamentación *de* la norma. Lo erróneo de esta concepción se aprecia claramente si se tiene en cuenta el aspecto comparativo de la fundamentación práctica: en el caso más simple, cuando contamos sólo con dos personas A y B, tanto A como B pueden disponer en una escala todos los sistemas normativos imaginables vigentes entre ambos. Pueden, pues, ponerse de acuerdo simplemente en un término medio. Mas —como subraya justamente Wolf— este acuerdo tiene el sentido de un compromiso, pero no de una fundamentación. Pues aquello que para uno estaría mejor fundado práctico-subjetivamente, estaría peor fundado para el otro.

No creo que se trate de un error de desarrollo, susceptible de una corrección inmanente. El error es más fundamental. Si la única instancia a la que se puede recurrir tras el ocaso de las verdades superiores es el interés recíproco de los individuos en el respeto de determinadas normas por parte de los demás, en el *problema de la fundamentación* no se puede ir más allá de lo que se ha designado en el primero de los dos puntos expuestos al comienzo del capítulo anterior («bueno para mí»). El paso de «bueno para mí» a «bueno para todos» es ciertamente obligado, pero no es el tránsito

a una fundamentación, sólo es necesario porque de otro modo los demás no colaborarían. Que yo tenga un motivo para someterme voluntariamente a aquel sistema normativo que es bueno para todos sólo es así en la medida en que éste es bueno para mí, sencillamente porque para mí es mejor que ningún sistema normativo o que uno que me perjudicase. De este modo no se puede fundar una moral, e incluso un sistema jurídico, de esta manera, sólo puede ser estipulado, pero no fundado.

Por tanto, la debilidad de mi intento de fundamentación puede superarse con la condición de retirar la premisa en que se basaba: la presuposición de que, con la supresión de las verdades superiores sólo puede recurrirse a los intereses de los individuos y ya no a una propiedad esencial. No hay motivo alguno para suponer que una concepción moral sólo apunta a una concepción de la identidad esencial de la persona si tiene un fundamento tradicional o metafísico; más bien, como se ha mostrado entretanto, esta conexión es constitutiva del sentido de la moralidad. Por consiguiente, sólo quedan dos posibilidades: una vez desechadas las verdades superiores, la moral no sólo no es fundable, sino que se convierte en una palabra vacía; o bien tendría que mostrarse que, aunque se desechen las verdades superiores, subsiste algo semejante a una propiedad esencial. Ésta tendría que ser constatada como hecho empírico.

Para responder a la cuestión de qué podría significar aún en este caso particular (en el que ya no está dada de antemano una verdad superior) estimar sin más a una persona, debo retomar una vez más la cuestión general —qué significa esto en general. Porque lo que hasta ahora he dicho sobre el particular resulta insuficiente. Por un lado, he tenido que hablar de una

«propiedad esencial», que consiste sustancialmente en cada caso en nuestra propia «autocomprensión», porque de otro modo no podría haber tenido en cuenta las concepciones tradicionalistas y metafísicas de uno mismo y de la moral. Pero precisamente esta concepción de la comprensión esencial de uno mismo no resultó ser concluyente (véase p. 167), y ahora podemos ver por qué no podía serlo: porque remite a una verdad superior. Esta debilidad no se da por tanto a expensas de mi explicación, sino de las propias concepciones morales tradicionales.

Pero por otra parte habíamos ya encontrado en otro contexto, mucho más simple, una conducta hacia uno mismo y hacia los demás en la que la persona es concebida en sí misma y no en cuanto le corresponde esta o aquella propiedad: en relación a los demás, el amor y la amistad (véase p. 164), en relación con uno mismo, la afirmación voluntaria del propio ser en el sentido de la voluntad de vivir (véase p. 154). Estas dos formas de conducta se corresponden. Lo que una es en primera persona, lo es la otra en segunda persona. En ambos casos se trata de una conducta relativa a una persona —en un caso la propia, y en el otro la de otro— *como tal* y no en un determinado aspecto, como el de si quiere esto o aquello o tiene esta o aquella propiedad. Con ello, ha encontrado una respuesta positiva la pregunta de qué se entiende por este «sí mismo» [*selbst*], una respuesta diferente de la tradicional: en términos ontológicos, no se trata ya de una propiedad en el sentido normal de la palabra (por tanto que tuviese que expresarse mediante un predicado), sino de la existencia. Esta determinación que en rigor no es una propiedad (que se expresa con un predicado) la denominaré de ahora en adelante una cuasipropiedad, pero seguiré hablando de una propiedad a propósito de una

cuasipropiedad, porque no creo sea de temer un malentendido.

Es un hecho empírico que tanto con respecto a nuestra propia existencia como a la existencia —la vida— de los demás podemos tener una relación volitiva (afirmadora o negadora), sea cual sea efectivamente esta relación. Este comportarse con uno mismo [*Sichverhalten*] —con mí mismo y con los demás en cuanto tales— está contenido en las concepciones tradicionalistas y metafísicas de aquello que somos «esencialmente», pero recubierto en cierta medida de una propiedad esencial. Si se desecha la verdad superior que estaba en la base de la creencia en la propiedad esencial, lo que queda no es simplemente un agregado de intereses, sino la cuasipropiedad, por así decir purificada, sin propiedad, del comportarse hacia sí mismo y los demás en cuanto [sujetos] que se comportan igualmente.

«Sí mismo» es una expresión de contraste. ¿De qué contraste se trata en el uso actual? La asociación que evoca en primerísimo lugar esta expresión, y que no es a la que nos referimos aquí, es la existente entre lo que cualquier cosa es respecto a su propiedad sustancial y sus restantes propiedades, consideradas accidentales. En el caso actual no se trata de propiedades, sino de objetos del querer: por un lado mis múltiples intereses, deseos, metas, y por otro mi ser, que me es ya dado de antemano, lo quiera o no, y con el que tengo que tener una cierta relación volitiva (afirmadora o negadora).

Este contraste puede aclararse quizá mediante un contraste análogo que hallamos en Kant: el contraste entre fines (metas) subjetivos y fines en sí. Los fines subjetivos son fines «que un ser racional se propone a

su antojo como efectos del propio obrar» (IV, 427),[8] mientras que un fin en sí es un fin autónomo (IV, 437), es decir, algo dado de antemano a la voluntad, que no está a su «disposición» (véase IV, 429). A menudo se ha dudado que este concepto de fin autónomo en general sea un concepto significativo. Naturalmente no lo es si se define «fin», como hace por lo demás el propio Kant, como cualquier cosa que debe ser realizada por la voluntad. Pero Kant puede eludir fácilmente esta objeción más bien superficial sustituyendo la expresión «fin en sí» o «fin objetivo» por «valor en sí» (o «valor intrínseco») (véase IV, 435). Que mi ser y el ser de los demás tenga un valor intrínseco es una afirmación que Kant sólo puede concebir por supuesto con ayuda de su concepto fuerte (metafísico) de razón. No podemos seguirle tan lejos, con nuestros débiles medios. Que, por otra parte, cuando habla de un fin en sí, Kant se aproxime mucho al hecho empírico de la propia existencia dada preliminarmente lo muestra en esta afirmación: «Así [como fin en sí mismo] el ser humano se representa necesariamente su propia existencia» (IV, 429).

En este pasaje, Kant prosigue diciendo: «pero así se representa también su existencia todo otro ser racional ... »; y —prosigue Kant— «por tanto se trata a la vez de un principio objetivo»; y el siguiente paso es la famosa segunda formulación del imperativo categórico: «Obra de modo que trates a la humanidad, tanto en tu persona como en la de cualquier otro, siempre al mismo tiempo como fin, y nunca meramente como medio».

8. Cito aquí a Kant según la edición de la Academia; el número romano indica el volumen, y el arábigo, la página. La obra de referencia es la *Fundamentación de la metafísica de las costumbres.*

Naturalmente, esta fundamentación es defectuosa. En un pasaje ulterior de Kant (IV, 437) se encuentra otra fundamentación de esta segunda formulación del imperativo categórico, una fundamentación que indudablemente era la determinante para él, pero que está muy lejos de mis reflexiones. Desearía adherirme más bien a esta primera argumentación de Kant. De forma análoga a su formulación «igual que para uno, así también para todos los demás», puedo decir también: igual que me comporto hacia mí mismo (mi ser), así (puedo constatar que) lo hacen también todos los demás. En este sentido —en cuanto me comporto volitivamente hacia mí mismo— soy como todos los demás, soy uno de ellos. En términos de la psicología del desarrollo, el niño conoce cuándo supera la fase del egocentrismo. Esta cuasipropiedad —ser uno de todos— está directamente ligada con la cuasipropiedad de comportarse con relación a sí mismo.

Esta propiedad puede insertarse ahora en el esquema general de la fundamentación de la moral que he expuesto en las páginas 172 y ss. A la pregunta de por qué tengo que limitar mi libertad según estas normas precisas, para ser objeto de afirmación por parte de todos, respondo así: porque estas normas tienen la propiedad de ser buenas para todos, y porque tú eres uno de todos. La moral así fundada es la moral del respeto recíproco, que ordena reconocerse (considerarse) prácticamente unos a otros como «fin en sí».

Naturalmente esta concepción del respeto, de la consideración de los demás (en el primero de los diversos significados expuestos en las páginas 150 y ss.) no deriva todavía simplemente del hecho de saberse uno entre todos. Éste sería el error que antes he atribuido a Kant. Tampoco la moral podría seguirse de la autoatribución de la propiedad esencial como tal, en el

caso de las demás concepciones morales. Es obvio que en el caso actual esto demostraría realmente demasiado: pues entonces, si alguien se comporta consigo mismo y se reconoce como uno entre todos, no podría menos de respetar a los demás, y esto no tendría el carácter de un mandato moral. También si se reconoce como uno entre todos, si se supera la fase del egocentrismo, obviamente existe todavía la posibilidad de una *lack of moral sense*. Igual que antes, la premisa inicial es el interés por un sentimiento del propio valor entendido intersubjetivamente, el interés de poder considerarse digno de estima. Con respecto a mi propiedad de ser uno entre todos, puedo ser estimado por todos (la sanción moral) sólo si realizo bien esta propiedad en el sentido de que respeto a los demás. Si la estima esencial recíproca —por tanto, aquella consideración que implica la conservación de la comunidad con los demás— no se refiere a otra propiedad esencial —que sólo podría estar fundada en una verdad superior— sólo se puede referir al hecho de entenderse prácticamente a sí mismo como uno entre todos, es decir, de respetarse mutuamente.

En mi opinión esto es lo que *queda* si la consideración o estima moral no tiene ulteriores puntos de referencia. Esta concepción según la cual la moral del respeto recíproco tiene un carácter residual implica que tal moral constituye además el núcleo de toda otra moral. A la moral del respeto recíproco se han superpuesto simplemente otros puntos de vista morales, igual que la cuasipropiedad empírica de atenerse a la propia vida y saberse uno entre todos ha estado siempre presente, simplemente recubierta por concepciones no empíricas adicionales relativas a lo que somos esencialmente (véase p. 179). Esta tesis, que la moral del respeto recíproco es una componente de

toda moral, tendría que ser probada empíricamente por un lado, y por otro es el resultado analítico de esta argumentación, porque el imperativo del respeto recíproco está analíticamente implicado en el concepto de la consideración recíproca de las personas como tales. Este mandato es por tanto el que queda cuando decaen todos los contenidos posibles, que estarían necesitados de una fundamentación propia. Resulta de la pura forma de la consideración recíproca en su valor de posición para la posibilidad de la autoafirmación.

Puedo ahora retomar, de otro modo, la tesis defendida en las «Tres lecciones», según la cual la moral de la consideración recíproca tiene un fundamento más fuerte que las demás concepciones morales por cuanto se basa en premisas más débiles. Ni una moral tradicionalista, basada en una verdad superior, ni la moral del respeto recíproco carecen de presupuestos. Pero la moral del respeto recíproco no recurre ya a un estándar aceptado y no ya fundamentable por su parte, que establezca cuál es la propiedad esencial «para nosotros» (véase p. 174). La propiedad a la que recurre (una cuasipropiedad) es por el contrario un hecho empírico. Pero de este hecho se deriva el mandato del respeto recíproco sólo para aquel que haya asumido en su voluntad el «debe» de la sanción moral. Este presupuesto, que se afirma la estima recíproca de las personas como tales, subyace en igual medida a todas las concepciones morales. Si alguien no quiere o puede entenderse así, no se le puede demostrar ninguna moral. La única argumentación que puede tener lugar aún está en mostrar que se pierde todo aquello que normalmente queremos si no se quiere o puede asentir a esto.

Paso ahora a examinar el problema —particularmente difícil según Ursula Wolf— de si —y cómo— puede fundarse la tesis de la universalidad de la moral

del respeto recíproco. Podría parecer que he resuelto de antemano este problema fundando la moral del respeto recíproco en el hecho de que uno se concibe a sí como «uno entre todos» (y por tanto que se concibe volitivamente con respecto a los propios fines como uno entre todos). Pero el concebirse a sí mismo como uno entre todos en este sentido deja abierta la cuestión del alcance del término «todos». Una referencia a *todos* (a una comunidad con *los* demás) definida de una forma siempre más precisa se da intrínsecamente en toda moral, porque también la estima moral contiene esta referencia a «todos»; pero en una moral interna tradicional este «todos nosotros» puede muy bien estar limitado a los miembros del grupo (o ser general, pero de modo que suscite una diferencia moralmente relevante entre aquellos a los que se pertenece y cuya estima importa, y los demás). A mí me parece claro que toda restricción semejante debe estar fundada; y no veo cómo pueda realizarse esta fundamentación sin recurrir a una verdad superior. Esta fundamentación suele atribuir además un valor propio [*Eigenwert*] al grupo con respecto al cual se define el «todos». Por ello, una semejante concepción moral limita el respeto de los individuos de dos formas: no se trata a los individuos como individuos, sino como individuos F, y esto presupone que, en general, no se trata sólo de los individuos como fines en sí, sino que se presuponen determinadas totalidades supraindividuales que deben ser los fines últimos. Es difícil ver cómo premisas tan fuertes puedan tener otra fundamentación que no sea religiosa. Donde no se dispone de un recurso semejante ni eventualmente de otras verdades superiores, donde la moral se reduce por tanto a la moral del respeto recíproco, no existe tampoco ya un motivo para limitar el «todos nosotros» a un grupo específico, o, en términos

más drásticos: no es ya posible fundamentar una restricción semejante. Entiendo, por tanto, lo siguiente: como el sentido del «respeto recíproco» incluye también, de antemano, una referencia a «todos» (a diferencia del amor o la amistad, el respeto, por su sentido mismo no se refiere a individuos particulares), la necesidad de fundamentación aparece entonces al revés de como presupone Wolf: lo que precisa una fundamentación particular no es nuestro deber de respetar a todos en general, sino el motivo por el que debemos respetar sólo a ciertos individuos con características especiales.

Ciertamente se puede volver a preguntar: ¿quién son todos en general? No veo cómo, sobre la base del enfoque desarrollado, esta pregunta pueda recibir una respuesta diferente de la siguiente: todos aquellos capaces de respeto recíproco. Surgen entonces las conocidas dificultades relativas a aquellos seres a los cuales no corresponde esta cuasipropiedad, pero sí una parte de ella, a saber la de tener sensaciones y eventualmente relacionarse volitivamente al propio ser: los niños pequeños, los fetos y los animales. En el contexto de estas «Retractaciones» no puedo adentrarme más en esta compleja problemática. Me tengo que limitar a indicar lo esencial: el hecho de que con respecto a la cuestión de si —y en qué medida— estamos obligados a respetar moralmente a los seres no nacidos y a los animales, no se llegue a un consenso, resulta inevitable. Sería ingenuo aceptar que aquí una determinada concepción fuese la verdadera. Si el respeto moral está ligado a una determinada cuasipropiedad, cuando los seres que no tienen esta propiedad pero sí una propiedad análoga (tener sensaciones en el caso de los animales, estar preparados para ... en el de los seres vivos no nacidos, durante el primer estadio presensorial),

parece ser inevitable consecuencia la duda sobre la medida de la aplicabilidad del respeto moral.[9]

Por lo tanto, en una moral sin una fundamentación tradicionalista, se refiere el «todos nosotros» —que pertenece a toda moral— a la comunidad de todos los seres humanos, e incluso de ese modo constituye esa comunidad («la humanidad»). A partir de aquí puede despejarse un equívoco que podría sugerir mi tesis de que la moral del respeto recíproco constituye además la base de toda otra moral. Hay que distinguir entre la moral del respeto recíproco y la moral del respeto recíproco universal. La moral del respeto recíproco universal no es obviamente parte de toda moral. Sólo

9. Remito a la atinada discusión del problema de la moral con respecto a los animales que se encuentra en el cap. V.3 del estudio de U. Wolf. Es verdad que, a partir de un enfoque como el suyo, en el que la reciprocidad no está inserta en el mismo concepto de la moral, esta problemática debe configurarse de forma diversa que en un enfoque que supone la reciprocidad. Esta última estaba incluida en la concepción que he desarrollado en las «Tres lecciones», mediante la concepción contractualista allí defendida, y en la concepción actual está incorporada mediante los conceptos de respeto y de estima recíprocos. Es verdad que de la posible estima recíproca que pertenece al sentido de la moral en general no se sigue incondicionalmente la necesidad de que los comportamientos moralmente sancionados tengan que distinguirse por su parte mediante la reciprocidad. Pero esta consecuencia es necesaria si tales comportamientos deben derivar «de la pura forma de la consideración recíproca en su valor de posición para la posibilidad de la autoafirmación» (p. 183), es decir, si los comportamientos moralmente exigidos no se pueden fundamentar con arreglo a verdades superiores. La transferencia de la obligación de respeto moral a seres que no pueden ser a su vez morales según mi concepción no es, por tanto, tan simple como parece en la posición de Wolf. La solución simple al problema la ha pagado Wolf al precio de sacrificar el carácter de obligación propio de las normas morales (y en él está arraigada la reciprocidad).

si el mandato del respeto no es limitado por otros factores, sólo fundamentables mediante verdades superiores, puede desplegar plenamente su sentido propio, y devenir así una moral del respeto universal. En este aspecto, la moral no fundada de modo tradicional no es ya un mero residuo, una componente de toda otra moral, sino que está contrapuesta a las demás concepciones morales, a causa de su consecuencia universalista. Este universalismo que pertenece en sí y por sí a la consecuencia interna del imperativo del respeto recíproco, en las concepciones morales y jurídicas tradicionales puede bloquearse de dos modos, externo e interno. Externamente, si no todos los seres humanos pertenecen a la comunidad moralmente relevante. Internamente, si aquellos que pertenecen a la comunidad moralmente relevante no deben ser respetados todos del mismo modo.

Por el contrario, esto significa que la moral del respeto recíproco —si no se recorta mediante ningún aditamento— se refiere no sólo a todos los seres humanos, sino además a todos del mismo modo. Los seres humanos no son por naturaleza iguales, pero en cuanto sujetos que podemos estimar recíprocamente no se distinguen unos de otros; por ello plantean una igual exigencia de respeto. Somos iguales como miembros de la comunidad moral. De nuevo no es la igualdad en el derecho al respeto —como ya en la universalidad— lo que precisa de una fundamentación particular, sino la desigualdad, y ésta —como la limitación en la extensión, sólo se consigue en nombre de una verdad superior.

Que deba estar moralmente justificado sólo aquel derecho que tiene en consideración a todos de la misma manera, no puede fundamentarse de forma contractual —como había supuesto en las «Tres lecciones»—

recurriendo a una supuesta igualdad previa y natural, sino sólo apelando a este concepto de igualdad moralmente fundado. En el seno de la tradición de la filosofía del derecho, el autor al que prefiero asociarme ahora es por tanto John Locke, que construyó el derecho recurriendo a un hipotético estado de naturaleza, que ya estaba definido mediante la ley moral —y que por su parte Locke simplemente ha presupuesto.

4. Nueva retractación: esbozo de una moral de la seriedad

Las reflexiones de las anteriores secciones muestran la dirección en la que según mi opinión actual debe entenderse la fundamentación de la moral del respeto universal. Los dos puntos de vista fundamentales eran: en primer lugar, también esta moral, como cualquier otra, se funda en la sanción específicamente moral de la estima de la persona como tal (lo que da su sentido al «debe» específicamente moral). En segundo lugar, como todas las morales ésta tiene que referirse a la autocomprensión esencial recíproca de las personas, a una propiedad o cuasipropiedad que se supone esencial y que, como éstas no pueden ser de otro modo objeto de estima, tiene que permitir una gradación según el criterio mejor/peor. Y esta moral tiene la característica particular de que aquello a lo que se refiere esta autocomprensión no puede basarse en una verdad superior, y por tanto tiene que ser necesariamente un hecho empírico.

De estas dos directrices fundamentales hay que distinguir los detalles de mi desarrollo efectivo. Con respecto a éste me siento menos seguro. He identificado la autocomprensión recíproca de las personas por el

hecho de que éstas se comportan todas ellas volitivamente en relación a su propia existencia (su vida) y que lo saben recíprocamente unos de otros. Algunos lectores considerarán muy dudoso este enfoque existencialista. Se trata, sin embargo, de un hecho desde mi punto de vista indudable, que en nuestra época Heidegger se limitó a redescubrir y que ya había sido reconocido por Aristóteles.[10] Para facilitar la comprensión de los lectores me he referido a una expresión kantiana posiblemente mucho más accesible, pero en rigor menos clara: que cada ser humano considere su propia existencia como fin en sí.

Mis dudas se refieren a la cuestión de en qué medida he conseguido fundamentar el mandato del respeto recíproco a partir de los tres factores siguientes: 1) la cuasipropiedad de tener que atenerse [*des Sich-verhalten-Müssens*] a la propia existencia dada de antemano; 2) la cuasipropiedad conexa de ser uno entre todos o de saberse como tal; y 3) la afirmación volitiva de la estima moral recíproca. Al lector atento no se le habrán escapado las incertidumbres de la acentuación en esta parte central de la argumentación.

En ciertos lugares podría parecer que es la posibilidad de la estima moral recíproca como tal lo que fundamenta el mandato del respeto recíproco. Esto es parcialmente exacto, en el sentido de que la posibilidad de la estima recíproca se constituye en la comunidad moral (la consideración recíproca en este sentido fuerte), para lo cual es condición necesaria aquella comunidad que se constituye en cuanto las personas se respetan recíprocamente (la comunidad de la consideración recíproca en sentido débil —lo que Kant ha

10. Véase Tugendhat, *Selbstbewußtsein und Selbstbestimmung*, pp. 176 y ss.

llamado «reino de los fines»). No obstante, el concepto de estima estaría vacío si no hubiese una propiedad de las personas como tales, que representase el objeto de la estima; por lo tanto, el tercer punto remite necesariamente al primero y al segundo.

Pero esto lleva a la cuestión de cuál de estas cuasipropiedades es la decisiva para la estima moral o de si éstas en realidad tienen que considerarse conjuntamente —como ya he sugerido. En primer lugar, podría parecer tentador renunciar a la cuasipropiedad «existencialista» y tomar como punto de partida la propiedad de ser uno entre todos o concebirse como tal. He propiciado esta tentación al decir que esta propiedad podía ser satisfecha bien (y por tanto mejor o peor) (p. 181). Pero si ahora consideramos esta segunda propiedad separada de la primera, en realidad no puede decirse que pueda ser realizada mejor o peor. Ya habíamos visto cómo la expresión «concebirse a sí mismo como uno entre todos» es ambigua. En el primer sentido —débil— significa que yo sé teóricamente que todos los demás se conciben a sí mismos como yo me concibo a mí. En este sentido débil, que es compatible con la *lack of moral sense*, no tiene sentido decir que se pueda realizar tal propiedad mejor o peor. Si por el contrario se entiende ya en el sentido de que se tiene prácticamente consideración por los demás, una vez más no puede decirse que se pueda cumplir mejor o peor porque entonces significa ya la bondad moral.

Por lo tanto (para que mi intento pueda ser sostenible) debería ser la primera cuasipropiedad, el atenerse a sí mismo, la que constituye un posible objeto de la estima moral. Ahora bien, a diferencia de la cuasipropiedad «ser uno entre todos», la cuasipropiedad del atenerse a uno mismo está constituida de hecho, desde el primer momento, por una polaridad

de dos posibilidades fundamentales (entre cuyos polos existen diferentes grados): puedo componérmelas [*mich aussetzen*] con mi existencia, puedo tomarla en serio o bien eludirla y «perderme» en los deseos y metas singulares que ocasionalmente se me presentan. Para estos polos —concebidos por él de forma peculiar—, Heidegger utilizó los términos «existencia auténtica» [*eigentlich*] e «inauténtica» [*uneigentlich*]. En mi obra *Selbstbewusstsein und Selbstbestimmung* [Autoconsciencia y autodeterminación] he intentado entender en relación con el «problema práctico» aquella «autenticidad» de Heidegger, y he sintetizado el resultado en el concepto de «responsabilidad» —existir responsablemente.[11]

Parece que estas dos posibilidades de atenerse a la propia vida no sólo se dan de hecho, sino que la existencia en el modo de la seriedad es evidentemente un objeto de estima, y que obviamente se trata de una estima de la persona *en cuanto tal*, porque es aquello que es ella misma, un ser que se atiene a sí mismo. Creo incluso que puede decirse: estimamos a las personas en su ser como personas (en tanto no hay en juego verdades superiores) sólo si tomamos en serio su existencia. Este «sólo si» podría parecer una exageración. ¿Acaso no estimamos también a las personas cuando respetan a los demás? Entonces, tenemos que llegar naturalmente a este resultado para que sea posible fundamentar una moral del respeto recíproco. Pero —se puede preguntar todavía—, ¿por qué debemos respetar a una persona como persona cuando ésta respeta a todos los demás? Esto sólo es sensato si el respeto de los demás forma parte de aquello que hace a una persona digna de estima en cuanto persona. Por

11. Véase *ibidem*, p. 295.

tanto, tendríamos que poder suponer que una persona, si se toma en serio —y es esto lo que la hace digna de estima en cuanto persona— toma también en serio a todas las demás. Probablemente de hecho es imposible respetar a los demás —en el sentido fuerte de que con ellos no sólo se cumplen aquellas prestaciones sustantivas que derivan del respeto (guardar las promesas, etc.; véase p. 152), sino además de tomarles en serio como personas—, si uno no se toma en serio a sí mismo. Al parecer vale también la implicación contraria. Es esto lo que quiere decir Fromm cuando dice que puedo amar a los demás sólo si me amo a mí mismo, y viceversa. Entonces el contraste decisivo entre existencia moral e inmoral no sería el existente entre altruismo y egoísmo, sino entre tomar en serio —a sí y a los demás— por un lado, y por otro el seguir los deseos y finalidades presentes en cada momento.

La concepción que aquí se perfila es, en sus resultados, comparable a la que ha expresado Kant en la segunda formulación del imperativo categórico, según la cual el mandato moral es: «obra de tal modo que trates a la humanidad, tanto en tu persona como en la de los demás, siempre al mismo tiempo como fin y nunca meramente como medio». En nuestro contexto actual, esta formulación es interesante porque el mandato moral se refiere por igual al modo en que me atengo a mí mismo y al modo en que me comporto con los demás.

Ya en la sección anterior he apuntado a la analogía que me parece existe entre mi distinción —comportarse volitivamente con cualesquiera fines por una parte y atenerse al propio ser como algo dado— y la distinción kantiana entre fines subjetivos y fines en sí. Allí había presentado la analogía de una forma que, cuando uno se atiene a su propio ser, esto corres-

ponde a la concepción kantiana de que uno se atiene a un fin en sí. Sólo que un atenerse al propio ser tiene lugar siempre de una manera u otra. De forma correspondiente, Kant podría decir: me atengo a mi propia existencia en cuanto fin en sí siempre de uno u otro modo. Pero el mandato moral dice: condúcete con ella de tal modo que la reconozcas [siempre] como fin en sí. Y esto corresponde a lo siguiente en la concepción ahora desarrollada: condúcete con respeto a tu vida en el modo de la seriedad.

La moral de la seriedad se distingue así de la moral del respeto recíproco universal en que es más global y en que contiene también un deber moral hacia uno mismo. La concepción habitual en la filosofía moral actual es que no existen deberes hacia uno mismo (ciertos filósofos llegan incluso a este resultado mediante la forma en que definen el concepto de moral, lo que obviamente es inadmisible; [12] véase p. 137). También yo he sostenido en las «Tres lecciones» la tesis de que los deberes hacia uno mismo no pueden ser fundados a falta de verdades superiores. Pero el deber hacia uno mismo, ahora resultante, atañe no a un contenido cualquiera sino sólo al cómo, al modo de atenerse a uno mismo.

La exégesis de Kant también está dominada por la opinión de que el intento kantiano de deducir deberes hacia uno mismo del propio imperativo categórico contiene un paralogismo. Esto es correcto con respecto a la primera formulación del imperativo categórico. Pero no es correcto con respecto a la segunda formulación. Contra la opinión del propio Kant, esta última es más general que la primera formulación, precisamente

12. Véase el trabajo clásico de C. H. Whiteley «On duties» para una presentación del tema. Recopilado en J. Feinberg, *Moral concepts*, Oxford, 1969. (Hay versión española.)

del mismo modo en que la moral de la seriedad es más universal que la moral del respeto. Cierto es que, si se consideran los deberes hacia uno mismo y los deberes para con los demás en relación con esta segunda formulación del imperativo categórico, como ha hecho Kant en la *Fundamentación de la metafísica de las costumbres*, con ello la consideración del «fin en sí» adopta diversos significados en uno y otro caso. Considero a los demás como un fin en sí, si obro de tal modo que éstos puedan «aprobar» el modo en que yo me comporto con respecto a ellos (IV, 430), por lo tanto si tengo en cuenta lo que ellos *desean*. Me atengo a mí mismo como fin en sí si acepto mi ser dado como algo que *no* está a disposición de mis deseos (IV, 429).

Podría objetarse a Kant que, nada más concretar la aplicación del singular concepto de fin en sí, éste parece escindirse en dos conceptos diferentes. Pero la asimetría que se aprecia aquí entre el deber para con los demás y el deber para con uno mismo es de carácter intrínseco. De cualquier modo, se revela la misma ambigüedad —si se trata de una ambigüedad— en el concepto de seriedad. Tomar en serio a otra persona significa respetar no sólo sus deseos, sino a ella misma. Me atengo seriamente a mí mismo si existo responsable y autónomamente en el ámbito práctico.

El propio Kant ha fundado con claridad esta simetría en la *Fundamentación de la metafísica de las costumbres*. A la pregunta: «¿Cuáles son los fines que son a la vez deberes?», responde: «la propia perfección y la felicidad ajena»; y explica: «no se puede cambiar la una por la otra». Pues la propia felicidad no puede ser deber: «Aquello que cada cual quiere inevitablemente de sí no entra en el concepto de deber». Pero no se puede tampoco proponer como fin la perfección del otro, «pues se contradice exigir ... que yo deba hacer algu-

na cosa que no pueda hacer nadie más que él mismo».[13] Deliberadamente, no he entrado a considerar las ideas sustantivas de Kant sobre el deber para con uno mismo; están ligadas a la concepción de un fin objetivo, y para nosotros hoy resultan menos plausibles. Pero en la frase antes citada su concepción se aproxima mucho a la de «asumirse a sí mismo en la autonomía» [*Sich-in-Autonomie-selbst-Übernehmens*] por mí desarrollada. Subsiste la asimetría en el tomar en serio a los demás y a sí mismo porque cada cual sólo puede realizar su propia autonomía o, en términos más generales: porque no podemos despojar a otro de su «atenerse a sí mismos» [*Sich-zu-sich-Verhaltens*]

Pero —podría preguntarse—, ¿acaso no contradice también el concepto de autonomía el que podamos exigirnos moralmente la autonomía recíprocamente unos a otros? Creo que no, y pienso además que la preocupación por la autonomía del otro no sólo atañe al amor y a la amistad, sino también al respeto. Así se recupera la unicidad del concepto de seriedad. No obstante, sólo he intentado apuntar el concepto, no explicarlo.

Los análisis de esta sección y las anteriores tienen un carácter tentativo. No pretendo disimular la tensión que subsiste entre las afirmaciones de ambos apartados con respecto a la fundamentación del respeto recíproco, sino ponerla de manifiesto. Por tanto, todavía no se ha fundado suficientemente la moral del respeto recíproco. Tampoco estoy seguro si es correcto hablar de un deber con respecto a uno mismo, pero creo que el problema debería ser tomado nuevamente en serio.

(1983)

13. Kant, *Metafísica de las costumbres*, edición de la Academia cit., vol. VI, pp. 385-386.

ÍNDICE

Prólogo 7

Observaciones sobre algunos aspectos metodológicos de Una teoría de la justicia *de Rawls* 15

Ética antigua y moderna (en el 80 aniversario de Gadamer) 39

Tres lecciones sobre problemas de la ética . 66
Observación preliminar 66
1. El acceso semántico a la moral . . . 69
2. ¿Se puede aprender moralmente a partir de la experiencia? 99
3. Moral y comunicación 122

Retractaciones 147
1. Discusión con Ursula Wolf 147
2. Consecuencias para el concepto de la moral 162
3. La moral del respeto recíproco . . . 174
4. Nueva retractación: esbozo de una moral de la seriedad 188

CRÍTICA/FILOSOFÍA
Directora: Victoria Camps

Títulos publicados:

Ludovico Geymonat
HISTORIA DE LA FILOSOFÍA
Y DE LA CIENCIA

1. Antigüedad y Edad Media
2. Del Renacimiento a la Ilustración
3. El pensamiento contemporáneo

A. J. Ayer
WITTGENSTEIN

Alasdair MacIntyre
TRAS LA VIRTUD

Ernst Tugendhat
PROBLEMAS DE LA ÉTICA

Ludwig Wittgenstein
INVESTIGACIONES FILOSÓFICAS

Victoria Camps, ed.
HISTORIA DE LA ÉTICA

1. De los griegos al Renacimiento
2. La ética moderna *(en prensa)*
3. La ética contemporánea *(en preparación)*

La pregunta más importante de la ética es la de su fundamentación. Una pregunta que, según Tugendhat, no tiene una respuesta definitiva ni válida para todos los tiempos. A su análisis, al intento de dar algunas respuestas y a la revisión de las mismas están dedicados los capítulos que componen estos *Problemas de la ética*. Las «Tres lecciones» que constituyen el núcleo central del libro van seguidas de las «Retractaciones» a las que ha llegado el autor tras la discusión de sus propias tesis con colegas tan distinguidos en el mundo filosófico como Jürgen Habermas y Ursula Wolf, asidua colaboradora de Tugendhat. Completan el texto un capítulo sobre *Una teoría de la justicia*, de John Rawls, y otro —homenaje a Gadamer— sobre «la vieja y la nueva ética». El pensamiento de Tugendhat es un claro exponente del encuentro de dos corrientes filosóficas contemporáneas: la fenomenología y el análisis del lenguaje. Representa la voluntad de pensar una ética comprensible para nuestro tiempo y nuestro mundo, nacida del diálogo y el debate con otras teorías del siglo XX. Los últimos escritos y manifestaciones diversas de Tugendhat a favor del pacifismo han extendido su prestigio más allá del ámbito académico y son una evidente muestra de su esfuerzo por entender la ética como un saber práctico.

ISBN: 84-7423-344-5